#2

2762

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Applicant:	§	
Arnold J. GOLDMAN et al	§	
Serial No.: 09/588,681	§	
Filed: June7, 2000	§	Group Art Unit:
For: A KNOWLEGE - ENGINEERING PROTOCOL SUITE	§	Attorney Docket No.:2104/2
	§	

OIPE SEP 2 5 2000 JC17 PATENT & TRADEMARK OFFICE

Examiner:

TRANSMITTAL LETTER

Commissioner of Patents and Trademarks
Washington, D. C. 20231

Dear Sir:

The following documents are enclosed:

(1) Certified copy of priority document (IL 132663).

RECEIVED
SEP 27 2000
TECH CENTER 2700

Respectfully submitted,

Mark M. Friedman
USPTO Registration No. 33,883

Date: September 19, 2000

מדינת ישראל

STATE OF ISRAEL

Ministry of Justice
Patent Office

משרד המשפטים
לשכת הפטנטים

OIPE JC17
SEP 25 2000
PATENT & TRADEMARK OFFICE

This is to certify that annexed hereto is a true copy of the documents as originally deposited with the patent application of which particulars are specified on the first page of the annex.

זאת לתעודה כי רצופים בזה העתקים נכונים של המסמכים שהופקדו לכתחילה עם הבקשה לפטנט לפי הפרטים הרשומים בעמוד הראשון של הנספח.

RECEIVED
SEP 27 2000
TECH CENTER 2700

This 29-08-2000 היום

לסרי
ממונה על הבוחנים
רשם הפטנטים
Commissioner of Patents

נתאשר
Certified

לשימוש הלשכה
For Office Use

132663/2	מספר: Number
31-Oct-99	תאריך: Date
	הוקדם/נדחה Ante/Post-dated

חוק הפטנטים, תשכ"ז - 1967
PATENT LAW, 5727 - 1967

בקשה לפטנט
Application For Patent

אני, (שם המבקש, מענו ולגבי גוף מאוגד - מקום התאגדותו)
I (Name and address of applicant, and in case of body corporate - place of incorporation)

Insyst Ltd.
P.O.Box 45179
Jerusalem 91450

אינסיסט בע"מ
ת.ד. 45179
ירושלים 91450

בעל אמצאה מכח הדין .... Operation of Law ששמה הוא
Owner, by virtue of .... of an invention the title of which is

(בעברית) (Hebrew) מערכת פרוטוקול להנדסת ידע

(באנגלית) (English) A KNOWLEDGE ENGINEERING PROTOCOL SUITE

מבקש בזאת כי ינתן לי עליה פטנט
hereby apply for a patent to be granted to me in respect thereof

• בקשת חלוקה - Application of Division	• בקשת פטנט מוסף - Application for Patent Addition	• דרישת דין קדימה Priority Claim		
מבקשת פטנט from Application	לבקשה/לפטנט to Patent/Appl.	מספר/סימן Number/Mark	תאריך Date	מדינת האיגוד Convention Country
No. .... מס'	No. .... מס'			
dated .... מיום	dated .... מיום			
יפוי כח: כללי P.O.A.: general/individual-attached/to be filed later- filed in case .... הוגש בעניין				
המען למסירת מכתבים בישראל Address for service in Israel DR. MARK FRIEDMAN LTD. BEIT SAMUELOFF 7 HAOMANIM STREET 67897 TEL AVIV				

חתימת המבקש
Signature of Applicant

היום 2 בחודש Aug שנת 2000
This 2 of Aug of the year 2000

לשימוש הלשכה
For Office Use

Ref.: 2104/4

A Knowledge-Engineering Protocol-Suite

מערכת פרוטוקול להנדסת ידע

ABSTRACT

A Knowledge-Engineering Protocol-Suite is presented generally includes methods, systems, and apparatus for search-space organizational validation; and appurtenances for use therewith. The protocol-suite includes a **search-space organizational validation method** for synergistically combining knowledge bases of disparate resolution data-sets, such as by actual or simulated integrating of lower resolution expert-experience based model-like templates to higher resolution empirical data-capture dense quantitative search-spaces. Furthermore, from alternative technological vantages, the suite relates to situations where this synergetic combining is beneficially accomplished, such as in control systems, command control systems, command control communications systems, computational apparatus associated with the aforesaid, and to quantitative modeling and measuring tools used therewith. The protocol-suite also includes facile **algorithmic tools** for use with the method and **a process-modeling computer** for use in a distributed asynchronous system of process modeling computers.

GENERAL FIELD OF THE INVENTION

The present invention generally relates to knowledge-engineering, to search-space organizational validation therein, and to protocol-suites for use therewith.

More specifically, the present invention relates to synergistically combining knowledge bases of disparate resolution data-sets, such as by actual or simulated integrating of lower resolution expert-experience based model-like templates to higher resolution empirical data-capture dense quantitative search-spaces.

Furthermore, given the inherent interdisciplinary nature of the present invention, from alternative technological vantages, the present invention may also be understood to relate to knowledge-engineering embodiments where this synergetic combining is beneficially accomplished, such as in control systems, command control systems, command control communications systems, computational apparatus associated with the aforesaid, and to quantitative modeling and measuring tools used therewith. Equivalently, the present invention may be understood to relate to domains in which this synergetic combining is applied, such as design and fabrication of integrated circuits, medical treatment modalities, social engineering models, corporate management enterprise systems, transactional modifications for financial business practices, or substantially any other organized modality of practice or information; technological, bio-physical, mercantile, social, etc.

GENERAL BACKGROUND OF THE INVENTION

In the fields of knowledge-engineering, database management, modeling, simulation, and expert systems, one common problem relates to forming valid optimization strategies over domains having constituent data-sets of assorted characters.

In this context, data-sets of assorted characters relates to data-sets that differ with respect to data structure complexity, to data resolution, to data quantification,

or to any combination thereof. Data structure complexity, data resolution, and data quantification may each relate to one-dimensional metrics or to multi-parametric characterizations.

In the context of the present document, data structure complexity, hereinafter "complexity", generally relates to local interconnectivity between a data element being characterized with respect to complexity and other data elements, and similarly global interconnectivity between any data-set, which includes this data element, and other data-sets. For example, a root node in a binary tree locally has two children branches of its own, and similarly may globally have many relationships that relate it to root nodes of other data structures.

In the context of the present document, data resolution, hereinafter "resolution", generally relates to an embedded relational concept wherein data-sets and proper data subsets are identified. The subset has a higher resolution than the superset, in that detailed data is placed in the subset while overview data is placed in the superset. For example, a superset may be a workflow overview organization, while subsets contain detailed charts of productivity measurements for each station in the workflow process.

In the context of the present document, data quantification, hereinafter "quantification", generally relates to a common sense notion of measurement precision. For example, in physics or chemistry it is common to measure a phenomenon to some known precision (e.g. velocity at mm/sec or pH to four decimal places), while in market surveys it is common to measure customer satisfaction using perhaps two to five select-only-one categories. While the average for a large number of surveyed customers may reach the same numerical precision as a physical measurement for a perhaps smaller number of samplings, nevertheless common sense still says that the physical measurement is a more realistic quantification than the survey result.

At the present juncture, it is necessary to appreciate that quantification disparities exist, and that known systems design methodologies encourage relating data-sets of like quantification while they discourage relating data-sets of

disparate quantification. Likewise, in a non-systems context, one could internally assign synthetic fractional quantification measures to semantic data-sets, and thereby presumably differentiate between their relative degrees of linguistic ambiguity, nomenclature variability, etc. However, synthetic fractional quantification measures used in a semantic environment would need to remain differentiated from quantification measures for their associated referents; at least so as to avoid semiotic symbol with referent confusions.

There are many examples of system-type problems related to forming valid optimization strategies over domains having constituent data-sets of assorted characters. According to one such example, there would be benefits if one could validly combine consumers' perceptions of fruit and vegetable quality with the agronomists' data capture universe; wherein is recorded precise measures of genetic makeup, growing conditions, biochemical variations, etc. According to another example, there would be benefits if one could validly combine demographic and actuarial databases with personal medical records and medical research data. Today, validly forming such strategies is a haphazard undertaking, of often-questionable objective value. More generally stated, there would be benefit if one could validly posit optimization strategies over domains having constituent data-sets of assorted characters; differing in complexity, resolution, and quantification.

Database management and knowledge-engineering represent a class of computer-implemented strategies for addressing such problems. Database management relates to organizational tools for establishing and maintaining data-sets of assorted character. For example, Boyce-Codd normal forms address tradeoff issues of efficiency and redundancy in very large purpose-specific data banks. However, database management does not address how to best benefit from knowledge that is held in these data banks.

Accordingly, there has arisen a discipline, currently called knowledge-engineering that attempts to generalize knowledge characterization strategies over heterogeneous domains having constituent data-sets of assorted characters; differing in complexity, resolution, and quantification. To date,

knowledge-engineering's most significant contribution has been the semantic search engine, which has subtle embodiment variations called search robots, search agents, data mining tools, etc. While search engines have proved to be very versatile tools for data-sets dominated by semantic content, they have not yet evolved into methodologies that provide meaningful linkages with data-sets having quantified characters. Thus the general need in the art remains to validly posit optimization strategies over domains having constituent data-sets of assorted characters; differing in complexity, resolution, and quantification.

A number of other classes of computer-implemented strategies are currently fashionable for addressing such problems. Examples of such strategies include modeling, expert systems, statistical process control, and neural networks. While each of these strategies has contributed some modest advance over its respective prior art, it is generally appreciated that these strategies are insufficiently modular to allow facile integration of new conceptualizations of ideas, which are brought into consideration by their implementation. Furthermore, the validity of the design process, which facilitates a computer implementation of any of these strategies, is often dependent on the level of genius of design team. Clearly, this is an inherent weakness, the alleviation of which would be of benefit in countless technological and econometric disciplines, especially if the method of alleviation is conceptually facile and straightforward for computerized implementation.

More specifically, a critical discussion of modeling, expert systems, statistical process control, and neural networks is forthcoming.

Modeling may be generally described as a low complexity topological graph describing node relations wherein each node corresponds to a data structure of empirical data. These nodes are homogeneously relating to a lower resolution and homogeneously relating to like quantification, while the associated data structures are disparately relating to higher resolution and to homogeneously like quantification within each data structure but not necessarily between data structures. The model is then used to simulate how the modeled system might react to a hypothetical perturbation of some of the empirical data.

Typically, modeling is applied in situations where there are many variables having complex interactions, especially where some of these interactions must be described using non-linear equations or using random variation functional components. Modeling is also applied in situations where visualizations, of the variables and their interactions, are believed to contribute to understanding aspects of the system being modeled.

Conceptually, the simplest models posit a pair-wise functional relationship between variables, such that each variable is a node of the topological graph and the pair-wise relationship describes the low complexity. The higher resolution data-sets then are used to describe an empirical manifold in the multi-dimensional space, as described by the pair-wise functionally orthogonal variables. Ordinary algebra, calculus, or statistics is then applied to simulate hypothetical empirical situations.

Conceptually, a more complex class of models posits multivariate functional relationship between assorted combinatorial groupings (n-tuples) of variables, wherein the aggregate of relationships join all of the variables into a single topological graph. Somewhat like the simpler models, higher resolution data-sets then are used to describe an empirical manifold for each relationship between the assorted combinatorial groupings of variables. Integrating a relational rule set with ordinary algebra, calculus, or statistics then allows hypothetical empirical situations to be simulated.

Conceptually, a most complex class of models posits embedding of either or both of the above described models within nodes of the more complex class of models. The designing and integrating of relational rules then becomes a cumbersome task that depends on the level of genius of design team, especially for computer implementations. Likewise, the classes of hypothetical empirical situations to be simulated are generally limited by the structure of the design.

In order to escape from this type of limitation, a tedious class of modeling tools called expert systems has been developed. Conceptually, expert systems shift the focus of the simulation from the empirical data manifolds to the designing and integrating of relational rules. Since it is presumed that the experts

have subsumed the empirical manifolds, simulating hypothetical empirical situations at the manifold level is replaced by simulating a higher complexity topological graph describing node relations. Expert systems then become a most complex class of models that are critically limited by the structure of their design. Methodologically, the only way to improve an expert system is by implementing a longitudinal study of interviewing experts and integrating their changes of mind and mood.

Another class of modeling tools, called process control models, has been developed. Here, the complexity of functional relationships between variables is grouped as a single node for each station in a process, and the topological graph of node relationships is according to the complexity of the process being modeled. Furthermore, each station in the process is internally amenable to any of the above modeling methodologies including expert systems, albeit as constrained by the inputs and outputs for each station. Independently, the overall process is likewise amenable to benefit from using any of the above modeling methodologies including expert systems, albeit as constrained by the topology of the process. Simply stated, process control focuses simulation and decision resources on a limited class of optimization hypotheses that are constrained by the topology of the process.

Process control models are chosen in circumstances where the overall process is pragmatically optimized by locally optimizing the process at each station. Furthermore, for most applications, process control focuses simulation and decision resources on a limited class of optimization hypotheses that are constrained by using the simplest modeling techniques for each station. For this reason, statistical process control tools, neural network tools, and similar tools have become popular, in that they can be facilely applied to any station, as if that station were isolated from factors at other stations.

In statistical process control (hereinafter SPC), gross statistically derived threshold-type limits are assigned individually for metrics associated with inputs or outputs at a station; wherein each of these metrics was considered in isolation,

in conceptually similar ways to that used in the simplest class of modeling and simulation.

For example, an SPC station may assemble two primitive components C1 and C2 together to form an aggregated component C3. Each of these components has statistically defined acceptable tolerance limits for at least one measurable aspect of the component; C1 (min, max), C2 (min, max), and C3 (min, max). The presumption is that if all C1 components are in the range C1 (min, max) and if all C2 components are in the range C2 (min, max), then all C3 components will be in the range C3 (min, max). Simply stated, using SPC tells us to set off an alarm and call a control process engineer whenever C3 components are measured to be out of the range C3 (min, max); and this actually happens even if C1 and C2 components were within their acceptable tolerance limits.

When out of specification C3 components are produced, the process control engineer first decides either to stop the process or to let the process continue. Typically the process is stopped when the result is potentially catastrophic, such as in nuclear power plant SPC or in chemical synthesis of essential therapeutic drugs. Otherwise, the process control engineer may elect to let the process continue, even though the resultant out of specification C3 components may be worth much less than in specification C3 components.

Regardless of the process control engineer's decision, there is a need in the art for a method of improving SPC. More specifically, there is a need in the art for automatic tools to aid the process control engineer in returning the process to producing C3-type components within acceptable tolerance limits.

One aspect of this standard SPC problem is that there is an accumulation of contingent degradation of tolerances, in a concatenation of specifications for a plurality of interdependent stations. Simply stated, when there is a plurality of independently defined specification limits, these specifications actually convolute at a higher resolution into a configuration where not every combination of input specification parameters yields an acceptable final station output result. Thus, there is a need in the art for a tool that allows SPC specifications to be convoluted at a higher combinatorial resolution.

Another way to appreciate this need is to consider SPC as a model of a multivariate functional relationship wherein an upper bound threshold manifold and a lower bound threshold manifold represent the solution limits for a predetermined volumetric region in an orthogonal solution space. Clearly, only in unusual circumstances, such as when the manifolds are parallel and also slice through the predetermined volumetric region in an absolutely orthogonal fashion, will the convolution of the SPC limits be equivalent for both low-resolution and high-resolution specifications. However, if the manifolds are parallel and also slice through the predetermined volumetric region in an absolutely orthogonal fashion, then virtually none of the variables in the domain of the multivariate functional relationships effect the results.

In neural networks, high-resolution empirical data is accumulated and correlated with low-resolution decision data, substantially in order to define limits like those that were defined in the SPC method. Neural networks are used in situations where setting specification threshold limits for inputs is excessively complex, often because input variables being measured are highly interdependent, and simultaneously where setting threshold limits for outputs is well understood or at least easy to define. Here too, there is a need in the art for a tool that contributes to defining acceptable tolerances for aspects of inputs to a neural network evaluated process, so as to beneficially improve metrics of productive throughput for that process.

Another way to appreciate this need is to consider a neural network as a model of a multivariate functional relationship wherein a very complex topological shape constitutes the solution limits for a predetermined volumetric region in an orthogonal solution space. While this may be correct, no additional understanding or progress may be derived from this solution. Therefore, when neural network are used, improvements and innovations of the process are conceptually inhibited.

In accordance with all of the aforesaid general background, there is a need in the art for a knowledge-engineering protocol-suite:

- to provide a unified frame of reference for the numerous aspects of knowledge-engineering;
- whereby new knowledge-engineering apparatus and appurtenances may be independently designed to integrate facilely with each other; and
- that substantially provides a framework through which existing knowledge-engineering products may be compared, functionally de-convoluted and seamlessly integrated to form large-scale knowledge-engineering systems.

Most professionals, working in knowledge-engineering, are familiar with the Open Systems Interconnect (OSI) reference model of the International Standards Organization (ISO). This well-known OSI model is a common point of reference for categorizing and describing network devices, protocols, and issues. Countless network devices are designed to operate at certain OSI protocol levels. Likewise, in today's ensemble of network protocols, virtually each of the known protocols can be mapped onto the OSI reference model. Accordingly, it would be of tremendous benefit if a knowledge-engineering protocol-suite could be provided that builds on this familiarity with the OSI model.

The (OSI) reference model offers a seven-layer model structure defining the "ideal" network communication architecture. This model allows communication software to be broken into modules. Each layer provides services needed by the next layer in a way that frees the upper layer from concern about how these services are provided. This simplifies the design of each layer.

With the emergence of open systems, the OSI model set rules that would allow different manufacturers to build products that would seamlessly interact. One of the key areas of importance is the interoperability of network technologies. As a result, this model was designed for the development of network protocols. Although no protocol has yet been developed using this

model, it has come to be accepted as a standard way of describing and categorizing existing protocols.

OSI conceptually puts names to the different tasks that a computer network has to fulfill. The ISO model defines seven layers, providing a logical grouping of the network functions. This model is good for teaching, and for planning the implementation of a computer network. Furthermore, dividing functionality in defined layers has the advantage that different parts of the network can be provided from different vendors and still work together.

When describing the different layers, one starts from the bottom and proceeds up through the upper layers. This is because some of the functionality and problems of the higher layers result from properties of the lower layers. The network stack used in the Internet illustrates the fact that a network is (usually) not implemented exactly as described in the OSI model. One protocol stack in use is referred to as the TCP/IP (Transfer Control Protocol/Internet Protocol) stack.

In order to appreciate today's network architectures and devices, it is important to understand the seven layers of the OSI model and their respective functions. The OSI reference model protocol layers, each with a unique function, are as follows:

• OSI Physical Layer (layer 1) is where the cable, connector and signaling specifications are defined. This layer provides mechanical, electrical, functional, and procedural means to activate and deactivate physical transmission connections between data-links. This layer is concerned with the encoding and decoding of digital bits (1s and 0s) between network interfaces. It is typically a function of the interface card, rather than a software utility.

• OSI Data-link Layer (layer 2) deals with getting data packets on and off the wire, error detection and correction and retransmission. This layer is generally broken into two sub-layers: The LLC (Logical Link Control) on the upper half, which does the error checking, and the MAC (Medium Access Control) on the

lower half, which deals with getting the data on and off the wire. This layer provides functional and procedural means for connectionless-mode transmission among networks. The data link layer is concerned with the transmission of packets from one network interface card to another, based on the physical address of the interface cards. Typical data link protocols are Token Ring and Ethernet. The device driver that comes with the network interface card typically enables these protocols. The device driver will be loaded in a specific order with the other protocol programs. The data link layer is a point-to-point protocol, much like an airline flight. If you have a direct flight, one plane can get you to your final destination. However, if you have a connecting flight, the plane gets you to your connection point, and another will get you from there to your destination, but its up to you to make the connection yourself. Bridges operate at this layer.

• OSI Network Layer (layer 3) makes certain that a packet sent from one device to another actually gets there in a reasonable period of time. Routing and flow controls are performed here. This is the lowest layer of the OSI model that can remain ignorant of the physical network. This layer provides a means of connectionless-mode transmission among transport entities. It makes transport entities independent of routing and relay considerations associated with connectionless-mode transmission. The network layer is concerned with the end-to-end delivery of messages. It operates on the basis of network addresses that are global in nature. Using the airline example, the network layer makes sure that all the connecting flights are made, so that you will actually arrive in your final destination. Network layer protocols include the IPX portion of the Netware IPX /SPX protocol and the IP portion of the TCP/IP protocol stack. Routers operate at this level.

• OSI Transport Layer (layer 4) makes sure the lower three layers are doing their job correctly, and provides a transparent, logical data stream between the end user and the network service being used. This is the lower layer that

provides local user services. This layer provides transparent data transfer between sessions and relieves them of concern about achieving reliable and cost effective data transfer. SUPER-UX supports Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). The transport layer is concerned with issues such as the safe, intact arrival of messages. It makes the receiver aware that it is going to receive a message, insures that it does get it, and can control the flow of the message if the receiver is getting it too fast, or re-transmit portions that arrive garbled. In our airline analogy, suppose you are flying your children to Grandma's house unaccompanied. The data link layer planes will make their flights. A small fee will insure that network layer ground attendants get your kids from one flight to their connection. The transport layer will call Grandma to let her know they are coming and what their luggage looks like, and will expect a call from Grandma when she has them safe and sound. Typical transport layer protocols are the SPX portion of Netware SPX /IPX and the TCP portion of TCP/IP.

• OSI Session Layer (layer 5) is where communications between applications across a network are controlled. Testing for out-of-sequence packets and handling two-way communication are handled here. This layer provides the services needed by protocols in the presentation layer to organize and synchronize their dialogue and manage data exchange. The session layer is the layer that manages all the activities of the layers below it. It does this by establishing what is called a virtual connection. Essentially a virtual connection is established when a transmitting station exchanges messages with the receiving station, and tells it to set up and maintain a communications link. This is similar to what happens when you log into the network. Once you have logged in, a connection is maintained throughout the course of your user session until you log out, even though you may not be accessing the network continuously.

• OSI Presentation Layer (layer 6) is where differences in data representation are dealt with. For example, UNIX-style line endings (CR only) might be converted to MS-DOS style (CRLF), or EBCIDIC to ASCII character sets. This layer manages the representation of the information that application layer protocols either communicate or reference during communication. The presentation layers function is to establish a common data format between communicating nodes. It is responsible for formatting the data in a way the receiving node can understand. It may also perform data translation between different data formats. Examples of data format differences include byte ordering (should it be read from left to right, or vice versa) and character set (ASCII characters or IBMs EBCDIC character set) as well as differences in numeric representation.

• OSI Application Layer (layer 7) is where the user applications software lies. Such issues as file access and transfer, virtual terminal emulation, inter-process communication and the like are handled here. This layer serves as the window between corresponding application processes that are exchanging information. The application layer provides the user-accessible services of the network. These services include such things as network file transfer and management, remote job initiation and control, virtual terminal sessions with attached hosts, electronic mail services, and network directory services.

This seven-layer OSI reference model has proved to be a great conceptual catalyst for today's rapid developments of network infrastructure apparatus and associated software systems. Recalling the definitions presented at the beginning of this general background section, specifically for "complexity", "resolution" and "quantification", there is a need in the art for models that can accommodate modeling domains that differ greatly with respect to "complexity", "resolution" and "quantification". More specifically, it would be of tremendous benefit if a single knowledge-engineering protocol-suite could not only be built on the existing familiarity with the OSI model but also be

facilely applied to disparate applications; such as those that differ greatly with respect to "complexity", "resolution" and "quantification".

The following technical articles and citations, patents, Internet accessible web-pages, and the like are thought to be useful for understanding the history of the art, the current state of the art, and the present needs and failings of the art. While it is presumed that the man of the art is already familiar with the substance conveyed by these items, others may find, in these items, concepts and descriptions that will advantageously supplement their appreciation of the present invention. Therefore, the citations given in this section **do not** constitute a disclosure for the man or the art, nor should they be considered as uniquely disclosing salient aspects of the prior art.

Expert Systems: Expert Systems – Design and Development – John Durkin; Prentice Hall International Inc. 1994, ISBN 0-13-348640-0, pp. 4-25.

Process Control: "Yield Analysis Software Solutions" – Pieter Burggraaf; Semiconductor International January 1996, pp. 79-85.

Statistical Process Control: Quality Control Handbook - Fourth Edition – J.M. Juran (Editor) McGraw-Hill Inc., 1988, 24.1-22 & 26.39-46.

Neural Networks: "An Introduction to computing with Neural Networks" – Richard P. Lippmann; IEEE ASSP Magazine April 1987, pp. 4-22.

GENERAL PRIOR ARTS TO THE INVENTION

The following technical articles and citations, patents, Internet accessible web-pages, and the like describe concepts, methods, systems, and apparatus useful for a better understanding of new, useful, or un-obvious aspects of the present invention; and implicitly therein for appreciating the innate inventive step

leading thereto. In juxtaposition to the citations presented in the general background section, the citations given in this section **do** constitute a disclosure for the man of the art, and should be considered as uniquely disclosing salient aspects of the prior art.

In this context, existing commercial product, which circumstantially derive from granted or pending patents, should be considered as exploiting the best enabling mode of the technology disclosed in those respective patents.

Commercial Products:

Knights Technology - Sunnyvale, Calif., U.S.A.
(www.knights.com)

Knights technology creates software systems that allow engineers to collect, correlate, analyze and report essential fab data and to try to determine sources of semiconductor yield loss and wafer defects. Knights has several programs and an encyclopedic trouble shooting guide. Knights gives its clients a very sophisticated but unintegrated tool kit. It can only leave the client dimly aware of the need of one smooth running global system that employs the variant pieces of software that are readily available today and would oversee production parameters and make adjustments, as necessary, automatically. Knights product also suffers from the built in limitation of being defect oriented. It is true they might be successful in correcting random yield loss but they will completely miss the cause of systematic yield loss.

ObjectSpace – 14850 Quorum Drive, Suite 500, Dallas Texas, U.S.A.
(www.objectspace.com)

ObjectSpace produces Advance Process Control (APC) software technology or system that enables Run-to-Run control and fault detection applications in the factory. The client would be better served if such software

were not limited to a one variable adjustment. ObjectSpace leaves the industry in need of a software technology that has a more global view of the fabrication process and incorporates "wafer history " into a more dynamic, self-correcting system.

Adventa – 3001 East Plano Parkway, Plano, Texas, U.S.A. (www.adventact.com)

Adventa produces a suite of products for the control and management of a production wafer fab. They are Control, Process, and Track WORKS respectively. Their product ProcessWORKS supports a model-based process control used in discrete manufacturing systems. Process models are used to calculate process settings on equipment for automatic recipe generation, based on desired process results. This makes for efficient production but limits correction to one locale, whereas the possible error or deviation maybe in anyone of many locales in the wafer history. Fixed formula negates the possibility of automatic self-adjustment and leaves the model in a static mode with limited overall optimization.

Domain Manufacturing Corporation – 63 South Ave., Burlington, Ma., U.S.A. (www.domainmfg.com)

Domain applies statistical measures of primary and secondary parameters or production data either as collected in real time during on line production or after completion of several production cycles. Their Pattern software detects and warns operators when abnormal process conditions occur. Off line Pattern's analysis capabilities can enable engineers to scan large volumes of data with the hope of identifying exceptional regions requiring further analysis and to assist engineers in identifying assignable causes. The aforementioned software does not provide possible solutions nor automatically expands its scope of analysis from the data it collates.

Semy Engineering, Inc. 2340 West Shangri La, Phoenix, Arizona, U.S.A. (www.SEMY.com)

Semy has a supervisory system and metrology tools that collects data from Advanced Run-to-Run Control closed loop systems. Based on the physical measurements derived from the metrology tools, user selected process parameters are automatically modified to keep the process centered. This application can be used to control a single step in the process using a feedback technique or it can automatically adjust a subsequent step based on the results of a previous step using a feed forward technique. The automatic adjustments are limited to the narrow parameters of the process recipes within specific limits established by the process engineer. This limits the any trouble-shooting to a local target not taking into consideration wafer history and leaves the user with a static model that cannot implement past data analysis into the present model.

HPL – San Jose Gateway Plaza, 2033 Gateway Place, San Jose, Calif., U.S.A. (www.hpl.com)

HPL offers a package of four standard Failure Analysis Navigation and Visualization solutions. Their complimentary software provides integrated access to all data that harbors yield loss cause information--product and process engineering and design data, in line fabrication data, test data and other data with the ability to add new data without changing application software. User interactive modes of operation with systematic correlation of information which "drill-down" to root causes of failures and yield limiters. When there is an alarm the engineers and design experts must come in and using a mining tool locate the defect and make the necessary adjustment. The system would be more effective if the model possessed a self-learning mode that would in future alarms situations be able to point to possible defect areas and suggest solutions, and in so doing would be able save valuable time and increase yield levels.

KLA – Tencor Corporation - 160 Rio Robles, San Jose, Calif., U.S.A. (WWW.KLA.COM)

KLA-Tencor manufactures a combination of hardware and software systems that have application in identifying and helping to reduce defects in integrated circuit fabrication. The KLA-Tencor yield management consultants must decide where and how much to sample. This methodology of fab yield evaluation paired with certain defect source analysis techniques hopefully may lead to a rapid isolation of a defect source. Once the fab parameters have been breached the defect becomes more readily observable, measured and located by the engineers if they can correctly interpret the software analysis and recommendations. They are saddled with the same limitation and narrowness of view as Knights in that they are defect oriented.

Triant Technologies Inc. – 20 Townsite Road, Nanaimo, BC Canada (www.triant.com)

Triant Technologies Inc.'s focus is on improving overall equipment effectiveness by providing solutions that increase equipment up time, minimizing the use of test wafers and to accrue useful data on process problem areas and reduce scrap. The companies monitoring components range from a data collection system to a real-time multivariate modeling system. The collected data is stored for both on and off line visualization. Both gross and subtle equipment faults are detected by the employment of set point and model-based monitoring and alarming. These technologies reduce false alarms and thus allow the process engineers to determine the source and cause of the fault in the fastest time possible. If the problem is not in the fabrication equipment then the speed in which the correction is made is no longer in the hands of Triant Technologies, but in the hands of other yield management and fault detection and analysis tools. Tools that Triant believes have reached their limit in terms of effectiveness. Tools

that employ models that are in the main defect driven, manual rather than automatic in their operational mode, and in the end static models lacking a self-learning ability and thus unable to suggest possible solutions once the alarm has been rung.

Yield Dynamics, Inc., Santa Clara, Calif., U.S.A. (www.ydyn.com)

Yield Dynamics markets a suite of seven products in yield analysis including data viewing, charting and analysis, wafer map data, data mining and advanced statistical tools. In the area of statistics they provide an option for multivariate analysis by adding a suite of advanced statistical algorithms to their product allowing for the viewing of many parameters simultaneously and hopefully uncovering relationships that standard univariate techniques are unable to capture due to their complicated interdependencies. The increase in analysis tools and the accumulation of more and more data cries out for an APC model that is more automated and dynamic; a self learning empirical model that provides a more holistic view of the fabrication process and incorporates an ability to point to the possible causes of the detected deviation and suggest solutions. The state of the art model is one that is capable of gathering increasingly larger and larger amounts of data that the engineers are forced to dig there way through with their "mining tools" in search of a possible solutions. A continued increase in data gathering with an apparent decrease in the ability to analyze it singles diminishing returns for the industry.

Essentially, the best available commercial products are not substantially better than the heretofore-available component tools, as described in the General Background of the Invention section. Accordingly, there remains a need in the art for more application independent tools, or for tools which will allow analysis down to the individual instantiation level, or for tools which will allow integration of empirically known units according to their actual interactions. More

particularly, there is a special need in the art for a method that will simultaneously facilitate progress for all of these diverse disjunctive needs.

ADVANTAGES, OBJECTS AND BENEFITS OF THE INVENTION

The knowledge-engineering protocol-suite of the present invention may be facilely applied to disparate applications; such as those that differ greatly with respect to "complexity", "resolution" and "quantification". In order to appreciate this as an advantage, a brief matrix of non-limiting examples will now be presented.

The sample illustrative examples in the General Background of the Invention section related to (firstly) combining consumers' perceptions of fruit and vegetable quality with the agronomists' data capture universe, and (secondly) to combining demographic and actuarial databases with personal medical records and medical research data. From a domain of knowledge-engineering problems, additional typical illustrative examples may be categorized according to nine discrete classification regions in the matrix. These nine regions are designated according to the parameters: "complexity and "quantification"; and therein (for each parameter) according to an initial subjective assessment categorization of High, middle, or low.

Initial Global Search-space **Complexity**

(High...Middle...Low) relates to:

- a measure of graph directed topology size such as total number of nodes, and
  - ranges of branching ratios therein such as inputs and outputs for a given node.

Characteristic Local Region **Quantify-ability**

(High...Middle...Low) relates to:

- a measure of nodes or of relationships in the graph directed topology characterizing
  - if it is the case that variables therein are numerically measured to a predetermined degree of precision (High or Middle) or
  - if it is the case that these variables are only designated to a logical categorization (Low).

While recognizing that the forthcoming matrix of typical example problems has been categorized arbitrarily, it should be appreciated that a shift in categorization would nevertheless be substantially transparent with respect to operations in the knowledge-engineering protocol-suite of the present invention.

The purpose of these forthcoming examples is to introduce a forum of discourse wherein the broad aspect benefits of the present invention may be further appreciated. By expanding the notions of Complexity and Quantify-ability, it will become evident that domain of problems on which the knowledge-engineering protocol-suite of the present invention operates is substantially broader than domains for methods of the prior art.

Accordingly, the method of the present invention will be understood as having greater utility than methods of the prior art. This improved utility is because the present method operates over a broader domain of problems. Furthermore, this improved utility is because the present method allows problems to be defined according to a plurality of perspectives. Ultimately, this improved utility is because the present method provides a convenient protocol suite compartmentalization for conceptualizing relationships between the perspectives and relevant empirical data sets, and therein provides facile tools for understanding and developing these relationships. In the context of the prior art, each of the forthcoming examples should be appreciated as representing a juxtaposition of perspectives with empirical data sets that heretofore demanded a large-scale custom-built software system.

Table of Contents for Matrix of Typical Example Problems

a) Sub Set Of Assembly Or Service Process

b) Gardening (or other common "How to" topics)

9) Complexity Measure Low & Quantify-Ability Measure Low

a) Customer Satisfaction Surveys

b) Voting Preference

NOTE: These typical problems are strictly non-limiting examples that have been selected so that the diverse domain of utility of the present invention may be broadly appreciated. Alternative illustrative examples might equally well be found in the many degrees of detail that are commonly employed when describing and operating in a large command control communications systems environment.

Brief Description for each of the Example Problems

(1a) Complexity Measure High & Quantify-Ability Measure High - Semi-Conductor Design and Fabrication Example:

This example relates to a network of events starting from a discussion about an initial design concept and concluding when a packaged semiconductor from a batch of substantially identically semiconductors is quality categorized by an end of process testing system. This network of events includes inter-relations between hundreds of thousands of related steps, sub-steps and variables. Sometimes this network of events includes upgrading CAD/CAM tools, apparatus in a fabrication facility, changing specifications to sub-contractors or suppliers, or even building a new fabrication facility.

(1b) Complexity Measure High & Quantify-Ability Measure High - Automotive Design and Fabrication Example:

This example relates to a network of events starting from a discussion about an initial design concept, continues with the eventual testing of a newly

manufactured vehicle, and concludes when all of the sales and maintenance reports are studied against the actual design and manufacture. Like the semiconductor example, this network of events includes inter-relations between hundreds of thousands of related steps, sub-steps and variables. Sometimes this network of events includes upgrading CAD/CAM tools, apparatus in a assembly plants, changing specifications to sub-contractors or suppliers, or even building a new assembly plant.

For classification purposes, the network of events of examples 1a or 1b should properly be represented as a model having a very large number of nodes wherein each node has complex inter-relationships (edges) with sometimes-large numbers of other nodes. In the context of such a model, many of the variables need to be recorded to high degrees of precision, both in the specifications and as measured at many stages in the fabrication (or manufacture) process.

A very large network of events of this type is related to, in the prior art, by dividing the network into many substantially independent sub-networks (often as long chains of nodes) and applying disparate tools to different sub-networks. For example, the design discussions may be managed using project management time tables and documentation version control indexes. Independently, sections of the fabrication (manufacture) may be managed using statistical process control techniques and design of experiment paradigms. Furthermore, the final results may be aggregated using gross measures of batch yield, customer satisfaction, and corporate profitability. Embodiments of the method of the present invention allows this fragmented management of a single network to be modeled and considered both as a global symbiotic milieu model and as an ensemble of synergetic separable local sub-models.

(2a) Complexity Measure High & Quantify-Ability Measure Middle - Multi-Scale/Resolution Models For Health Optimization Example:

This example relates to the wealth of health related data that exists, and to the seemingly insurmountable problem of how best to integrate this data so as to accrue its highest benefits. This example is a categorized elaboration of one of the introductory examples; combining demographic and actuarial databases with personal medical records and medical research data.

Today, life-relevant data collection is a decentralized parallel process of capturing aspects of public health statistics, actuarial records, medical research, individuals' health profiles, and their respective longitudinal accumulations. While it is generally appreciated that the topological complexity of a hypothetical graph of nodes and relationships integrating data from these numerous sources is at least as complex as examples (1a) and (1b), there is no prior art method, for integrating the disparately quantified data-sets included therein, that can provide a sufficiently beneficial result from the integration.

Embodiments of the method of the present invention provide modalities whereby an individual may be related to a plurality of data-sets that describe him, or his ancestors, or persons having a profile-resemblance to him, or groups to which one of the aforesaid belong. These embodiments may portray this individual in his relations to these other individuals and groups. Furthermore, these embodiments may then quantitatively posit and quantitatively test hypotheses about the individual or about groups of individuals. This may provide many new opportunities for superior results in managing health care for individuals, in managing public health policy, in improving actuarial table precision, etc.

(2b) Complexity Measure High & Quantify-Ability Measure Middle - Ordinary" Medical Diagnosis & Treatment Example:

This example relates to the process of improving the health of a patient's health, regardless of whether the patient is sick or healthy. The first stage of this process includes combining subjective observations by a patient, objective observations by that patient's medical-service professionals, and quantitative clinical pathology metrics for the patient. The second stage of this process

includes a definitive analysis of the patient's health, a prognosis for that patient, a strategy to improve the patient's health, and often a follow-up procedure that iterates another pairing of these first and second stages.

In this example, there are virtually endless potential categories and combinations of categories for the observations and metrics of the first stage. Simultaneously, these categories include significant subjective qualitative data, objective low precision data, and objective high precision data. Classically, the analysis first focuses on how to reduce the complexity of the data, substantially by pruning away as much of the general medical profile data about the patient as may be eliminated without risking any inadvertent eliminating of an optimal strategy option. Within this pruned data graph topology, there is a need to apply the disparate precision data and to then focus the results into a custom health strategy for the patient. Briefly stated, this is the art of medicine.

When applied to this problem, embodiments of the method of the present invention may be configured to resemble an overly conservative physician who performs the pruning of the data graph topology and the applying of the disparate precession data. These method-enabled pruning and filtering operations should save the skilled physician considerable time when positing a patient specific health improvement strategy. Alternatively, these method-enabled operations should permit the physician to expend greater consideration on the actual object of the process, achieving a best possible health improvement for a specific patient.

(3a) Complexity Measure High & Quantify-Ability Measure Low - Triage (Medical Emergency Classification & Prioritization Of Casualties) Example:

This example relates to a classic operations research problem wherein all available substantially external data about a casualty is juxtaposed against available medical resources (facilities, supply, personnel, etc) in order to classify the casualty as destined for initial treatment: immediately, as soon as possible after those classified immediately are treated, or eventuality.

This problem essentially attempts to transform a topologically complex set of interrelated physiological observations into a simple decision result. Existing triage models, while attempting to consider these interrelated physiological observations "scientifically", usually focus on the actual decision that needs to be made; given the limited medical resources of the actual situation. Therefore, seemingly external considerations (such as medical treatment success statistics, short term and long term costs, and expected resultant life "quality") often dominate in choosing a triage decision model.

Embodiments of the method of the present invention may be used to integrate physiological data and actual casualty data with existing triage models in order to test if any of these models objectively deliver the results that they expect to deliver. Alternatively, embodiments of the present invention may be used to derive new triage models, test-simulate them, and compare them to known field-tested triage models.

(3b) Complexity Measure High & Quantify-Ability Measure Low - Occupational Measurement And Tracking Example:

This example relates to well-appreciated problem of comparatively evaluating and proportionately compensating employees. This problem is further complicated by a desire of the employer to improve the productivity of his employees individually and to develop an optimally integrated organization; presumably in order to better compete with like employers. Job categories, skill categories, and metrics of productivity are often simplistic and subjective; depending on non-standardized evaluations by supervisors or coworkers. Nevertheless, in large corporate organizations, a mapping of skills, tasks, productivity, and workflow will quickly grow into a highly complex model; for which the dynamics of improvement are often axiomatic rather than scientific.

Embodiments of the method of the present invention may be applied to organize data about workflow, skills, evaluation, etc. Thereafter, the present method may be used to test these fleeting dogmatic axioms of management, to

posit more individualized alternatives, and to quantitatively validate these alternatives.

Occupational Measurement and Tracking

Many human resources functions such as selection, recruiting, placement, or career development call for the use of behavioral measurement. Jobs and work settings have their own individual qualities and characteristics. Each job and work setting can be viewed as the shadow or outline of the particular person that would best fill it. Behavioral measurement samples attitudes, skills, and psychological traits that are important for placement in particular work settings, specific positions, or occupations. Computer software has been developed for monitoring home and work life activities in terms of level of sophistication as well as a tool for measuring emotional availability and interpersonal sophistication which are which can be used in adult training situations in both instructing and assessing progress of trainees. There are number of instruments that contribute to the hiring process and screening instruments i.e. Literacy and Numeracy Test are used by a large manufacturer to direct an in-house training/upgrading program. A Career Planning/Competency Model encourages individuals to explore and gather information, which enables them to synthesize, gain competencies, make decisions, set goals and take action. Each individual varies in their progress through each of these stages for many reasons. Some advance rapidly through each or all of the stages while others progress more slowly. It is here in this major gap between all the models, and tests and measurements where the individual drops through or slips by. The test-models are no more and maybe less successful in the selection of suitable candidates for a particular slot than a competent managers intuition and experience. The present invention could incorporate into a model this intuition and experience and by doing so improve the "thinking" of the present software that takes into account many of the multitude of variables (perhaps emphasizing the wrong ones) that are supposed to in theory make one person more successful at a particular job over another person but in reality remains rather vague and inaccurate.

Complexity Measure Middle & Quantify-Ability Measure High

Macro-Economics and Experimental Physics

Production and consumption. What is produced and its cost. Who consumes and what they pay. Weather, war and peace and their effect on the marketplace can be measured down to the penny and pork belly at the end of any given period!

Graphs and models representing scarcity, opportunity costs, production possibilities, supply and demand, output, national income, budgets, deficits, the national debt, inflation, unemployment, foreign exchange, balance of payments, supply side economics .are some of the many variables that go into making up a fiscal policy which is the static model that nations use to navigate the very dynamic seas of international economics.

The Federal Reserve promotes international monetary cooperation, stability, and orderly exchange arrangements to foster economic growth and high levels of employment as well as to provide temporary financial assistance to countries under adequate safeguards to help ease balance of payments adjustment. This "fudge" will hopefully validate the fiscal policy by trying to limit "outside" variables. More likely there will be a completely new fiscal policy the following year.

Embodiments of the present invention may be applied to allow for a greater understanding of day to day changes or even hour to hour sea changes with suggestions of their meaning and significance. In that the present invention is multivariate and dynamic (self-learning) it has the potential to validate independent values and make the necessary adjustments and thus be more robust. Adjustments, feedback and feedforward (thus benefiting from the high quantifi-ablity end) could be made in real time instead of at the end of set periods and not have to be locked into a model's preset parameters. Thus the present invention would not only allow for real time monitoring but also be able to make predictions and suggest possible adjustments or corrections.

Experimental Physics

The research in Experimental Physics covers a wide range of phenomena, from the subatomic scale over atomic, molecularand condensed matter physics, to environmental physics and interdisciplinary astronomy including many examples of both fundamental and applied physics. In the study of electronic structure of solid materials, including their surfaces and interfaces, photoemission, inverse photoemission and other surface sensitive techniques are used. The researcher is limited to the tools in his kit. Based on previous experiments and their models he has a good idea of what he would like to confirm, prove or discover and with this in mind he sets up his experiment. The present invention would give him the added advantage of narrowing down a wide but not overwhelming array of variables. To discover from previous experiment-models the common modalities and their hierarchy of importance through a dynamic feedback and feedfoward process. For example the main tool in this case is an angle-resolved photoelectron spectroscopy in the UV range. Detailed and highly precise information can be obtained about valence states in the volume, surface states and resonances, chemical shifts of core levels and more with this device. The present invention utilizing the information of past models and the precise results of recent experiments would first remove irrelevant variables and add previously neglected ones based on the self-learning enablement and dynaism of the of the newly generated model. This enhancement would lend for a greater ability to predict, with ever-increasing accuracy, the results of future experiments and would be able to eliminate unnecessary ones and in so doing would save valuable time and considerable funds. Further the robustness of the present invention: its ability to self -correct and ignore variables that are irrelevant, could allow for the transformation of an experimental model into a practical one given its enhanced ability to project and predict future results. In this case it would show, how most efficently, this particular information could be applied in the use and improvement of surfaces and interfaces of semiconducting systems.

Complexity Measure Middle & Quantify-Ability Measure Middle

Psychiatric Behavior Intervention

Early intervention plans have multiple uses and have been developed from a technical and research base, with the hope of short circuiting future adult anti-social behavior. Level of care assessment and descriptions are objective, based on a child's age, mental health and behavior. It is a uniform system based on research; Q&A forms, interviews and statistics. Working with demographics, socio-ecomic profiles and various psychological and intelligence tests one can possibly get a vague outline of an individual's mental health at any given time. Objective measures are used as as required by health management policies and program evaluation criteria. Statistical analysis programs, evaluation, and report generation, including methods and procedures, have to meet the requirements of research-based protocols. Also, more than ever, it is being recognized that "early intervention programs" not only benefit children when they need it the most, but also provide a valuable tool for "cost cutting strategies." This is most evident in the improvement of school attendance, decreased school behavioral referrals and county foster care placements. County agencies, which provide effective early intervention programs, save up to thousands of dollars per day in foster care placement. This is not to mention the need to address the severe problems with the growing violence school children are exhibiting. Yet the emphasis on the financial reward of early interventions linked with the isolation of the child in the intervention procedure might have dire consequences in the long run. Today the limits of intervention or behavior modification can only be measured, by short-term observation and repeated testing and monitoring. With the rapid changes in society many of the tools for perceiving the psychological person becomes outdated well before they can be replaced. This is also true of stastical analysis, based on past social history, used in an effort to predict future social or anti-social behavior. The present invention would enable a model to self-correct and remain relevant in its interpretation of the social patterns that are constantly evolving as well as allowing for greater individualization. The intuitve fear of

aggressive behavior and violence on the part of the medical care community might give more wieght to certain variables than they deserve and at the same time overlook others that play a greater role than previously realized.

Psychophysics Experiments—

Psychophysics is the psychological study of relationships between physical stimuli and sensory response; in this case it's vision and perception. This involves the collecting of "effects" explained by a theory of perception. The Craik-O'Brien-Cornsweet effect (COCE) involves two adjacent figures that are identical in luminance profile (i.e., in distributions of absolute measurements of reflected light) but differ in brightness (i.e., in the subjective perception of lightness and darkness). The two regions are identical in terms of the objective property of luminance profile, but one looks darker than the other. The difference in brightness between rectangles depends upon the difference in luminance at the borders. Effects such as COCE present problems, which it is, the business of theoretical work in vision to solve. Any viable model of the human visual system is constrained in the sense that their output should correspond to the percept when their input corresponds to the stimulus. Again, the datum presented by this effect and to be explained by a theory of vision is a relation between phenomenological properties (how things look) and physical properties (how the patches reflect light). The reason it counts as a psychological effect is because the curve describing the brightness profile of the percept does not match the curve describing the luminance profile of the stimulus. What is reported in this effect is that one patch looks brighter than the other, even though there is no difference in luminance. And it is hard to see how "looking brighter" can be anything other than a comparison in terms of phenomenological properties. (It cannot, for example, be paraphrased in terms of detecting differences in luminance, because in this case those differences do not exist.) There is no other way to get at brightness as a datum other than by examining your own percepts or accepting other people's reports of their percepts. In this case the present invention would be able to expand

the scope of the experiment by bringing in and comparing variables from many experiments on perception, based on different social and cultural groupings, and how these particular brightness curves compare to those describing luminance in the original experiment. There, of course, might be missing variables in the measuring of luminance as well, which might be the hidden factor for why one rectangle appears brighter when there is no apparent measurable difference. A model enhanced be the present invention would take into account that the measure of particular luminance may posses undetected differences for how is it possible that the eyes and mind of one person being so different than that of another perceives the same difference in brightness, being that it is a psychological effect and the psychology of one is so different from another.

Complexity Measure Middle & Quantify-Ability Measure Low

Market Research

Traditional consumer industries are in transition. The Internet is transforming the competitive landscape and business models of traditional consumer markets. Market Modules, one of the new models claims: "...we allow clients to take the pulse of their specific industry and its place in the consumer Internet economy...these analyst-supported services are a supplement to sector-specific market sizing, 'best practices' profiles of key competitors, proprietary consumer data, and deal-by-deal analysis."

Online in the Media Mix says, "Integrating online and offline advertising campaigns isn't just a matter of slapping a Web site address onto a magazine or a television ad. True campaign integration involves creating advertising across all media that delivers a similar message, draws on the same creative look and feel, and aims to build brand and sales over the long term." Another market researcher and Ad agency, claims that, "Banner advertising is the proven, efficient way to attract new customers on dozens of high quality, subject

focused internet sites, with a single buy. Your customers are out there surfing the web and we can deliver them directly to you."

They might not know how exactly to sell the clients product but they can sell advertising just by claiming they have the "know how." They certainly know several of the variables that motivate people to buy. The reasons people buy can run the gamut from the psychological (i.e.; impulse, status, fantasy projection, self-worth, lack of self worth,) to the practical (i.e.; time saving, life saving, labor saving etc.) Through a variety of testing and experience the ad people have arrived a few basic sure-fire principals the main ones being repetition, repetition, repetition. And that takes money, money, money. So they throw together all the ingredients this way and that way and some times they get a cake. What the product is could be the least important selling point!

The present invention would allow for a better long-term view of past and project means of advertising. Internet advertising has more in common with billboard highway advertising than with let's say television. What was successful fifty years ago, what type of product and what type of consumer. This could be integrated with psycho – physics models that try to understand how the eye perceives the stimuli. The model could have both a feedback as well as a feed-forward adjustment that could take in new information such as economic changes and how they are effecting consumer habits.

Complexity Measure Low & Quantify-Ability Measure High

Engine Control

The TEC (Total Engine Control) is the most advanced engine management system available in world today. It combines state-of-the-art fuel injection with the industry leading Direct Ignition System, and replaces the existing carburetor or fuel injection system and ignition distributor and coil. The use of advanced, digital computer based technology offers the absolute best power-delivery system, smoothest "drive-ability" and lowest fuel

consumption possible. The TEC-I series of engine control units consist of a Direct Fire Unit that holds the coils, and a TEC controller which holds the injector drive circuits and control logic. This configuration is ideal for extremely powerful engines with multiple injectors at each cylinder. These hand built and custom configured systems are purpose built special order items. The model dictates that specific engine specs will deliver a certain output. When these expectations are not meet the engineers must investigate where the fault lies. What is needed is a model that measures the expected output of standard variables in the system against its actual output rather than a model that only predicts overall system output. The present invention would be able to establish a more global model that would help to increase the optimization of the whole system. At that point more esoteric and overlooked variables could begin to be added to the model in its empirical self-learning capacity.

One might equally well describe an example of Cow Life Cycle in terms similar to any other management or process description. Specifically, one monitors genetic makeup, health, and nutrition with an aim of optimization milk or meat production; clearly on an individual animal basis rather than on any larger grouping.

Likewise any Sub Set of a larger process control problem, relating Assembly or Service Process would benefit from improved process control.

Other examples might include common Gardening, or any of a myriad of processes which are often addressed in "How To" books or articles.

There is a benefit in using empirical results from a population to replace testing on global empirical data-sets for populations with testing on individual instantiations and thereby to direct optimization and decision processes into higher resolution empirical data.

Technological Need Issues: There is a need in the art for facile tools that will improve the efficiency or durability of individual physical machines, each of which actually differs from the theoretical machine upon which it was originally based. The knowledge-engineering protocol-suite of the present invention provides the context to define, develop, integrate, and test such tools. More specifically, the present invention provides such tools embodied as methods, systems, and apparatus for search-space organizational validation; and as other appurtenances developed for use with the knowledge-engineering protocol-suite.

Ergonomic Need Issues: There is a need in the art for facile tools that will improve the efficiency or durability of individual physical organisms (e.g. in human medical treatment or in veterinary applications), each of which actually differs from a class or sub-class of theoretical organisms that has been collectively studied and tested. The knowledge-engineering protocol-suite of the present invention provides a facility for developing such tools.

Economic Need Issues: There is a need in the art for facile tools that will improve the efficiency or durability of individual instantiations of systems integration, each of which actually differs from the theoretical designs upon which it was originally based. Simply stated, regardless of the complexity of any specific model, and regardless of the degree of technical expertise that may be necessary to postulate improvements for such model, there is a need in the art to accurately present a present situation (including its options) in a format that a non-technical manager can appreciate. This need is most acute when the model relates to unique, expensive, or very large individual system instantiations. The knowledge-engineering protocol-suite of the present invention is configurable for integrating considerations when relating to such modeling situations.

NOTICES

The reader should appreciate that a reference to an existing commercial product, which circumstantially derives from granted or pending patents, should be considered as a reference to the present best enabling mode of the technology disclosed in those patents.

Furthermore, numbers, alphabetic characters, and roman symbols are designated in the following sections for convenience of explanations only, and should by no means be regarded as imposing particular order on any method steps. Likewise, the present invention will forthwith be described with a certain degree of particularity, however those versed in the art will readily appreciate that various modifications and alterations may be carried out without departing from either the spirit or scope, as hereinafter claimed.

In describing the present invention, explanations are presented in light of currently accepted scientific theories and models. Such theories and models are subject to changes, both adiabatic and radical. Often these changes occur because representations for fundamental component elements are innovated, because new transformations between these elements are conceived, or because new interpretations arise for these elements or for their transformations. Therefore, it is important to note that the present invention relates to specific technological actualization in embodiments. Accordingly, theory or model dependent explanations herein, related to these embodiments, are presented for the purpose of teaching, the current man of the art or the current team of the art, how these embodiments may be substantially realized in practice. Alternative or equivalent explanations for these embodiments may neither deny nor alter their realization.

GLOSSARY of Terms Used in Documenting the Prototype

The **Empirical Controller** (**E-C** concept consists of several components described in the sequel. The qualitative component of the invention that integrates physical knowledge and logical understandings into a homogenetic knowledge structure is called the **Knowledge Tree** (**K-T.** The Knowledge Tree is displayed graphically as a directed network with nodes, which are called

Interconnection Cells. These cells express the local relationship between input and output process parameter measurements. The **POEM** algorithmic approach is applied to obtain (from process measurement data) the precise quantitative relationship at each cell. Each Interconnection Cell is converted to an **Interconnection Model** or **Model** in short. The Model contains the quantitative relationships between input and output. The Knowledge Tree together with this quantitative layer yields the **Empirical Model**. The Empirical Model serves as a multivariable characterization of the process being described, and can be used to predict and control process behavior. The component of the invention that sits on top of the Empirical Model, and converts human determined operational objectives into system useable form is called the **Automated Decision-Maker (ADM** alternatively **ADAM)**. The ADM operates and analyzes the Empirical Model to determine solutions that best meet the specified objectives and constraints. The entire three-tier structure consisting of the ADM, the Empirical Model, and the Knowledge Tree is refereed to as the **Empirical Controller** The Empirical Controller is a generic learning and thinking system, which performs **Empirical Control.**

The ***product*** that contains the E-C technology and is the realization of the Automated Decision-Maker is referred to as the **Adam (Automated Decision-Maker**). Adam serves as a global Decision-Maker tool, encompassing the entire process. The E-C technology when embodied in a product and used for intermediate process control of work groups or equipment clusters is referred to as the **Eden (Empirical Decision Enabling Network).** The E-C technology when embodied in a product and used for troubleshooting, optimization and control at the processing equipment or measuring tool level is referred to as the **Eve (Equipment Variable Evaluator).**

GENERAL OVERVIEW AND SUMMARY OF THE INVENTION

The present invention relates to **a knowledge-engineering protocol-suite** for facilitating open systems interconnection transactions in a multi-layer knowledge-engineering reference model substantially having

Layer 1 - a physical layer for interfacing with apparatus;

Layer 2 - a data-link layer for facilitating data-communications within any of these Layers 1-7 or between any plurality of these Layers 1-7;

Layer 3 - a network layer for maintaining transactional access to data ensembles;

Layer 4 - a transport layer for organizing and maintaining token correspondences and adjacency lists wherein are represented network layer relationships between the data sets or between elements in the data sets;

Layer 5 - a session layer for validating the transport layer represented relationships and for simulating alternative transport layer relationships;

Layer 6 - a presentation layer for designing and executing experimental session layer simulations, evaluations thereof and modifications thereto; and

Layer 7 - an application layer for prioritizing n-tuple strategy dynamics of presentation layer transactions;

wherein the knowledge-engineering protocol-suite includes:

A) either a structured system having

I) at least one process-management computer with a program for relating Layers 1-3,

II) at least one computer embodying a search-space organizational validation method program for relating Layers 3-5, and

III) at least one knowledge-engineering workstation with a program for relating Layers 5-7;

B) or equivalently a distributed asynchronous system of process-modeling computers with programs for relating Layers 1-7.

Generally, the present invention relates to programs for facilitating open systems interconnection transactions in the frame of reference of a multi-layer knowledge-engineering reference model using **a knowledge-engineering**

protocol-suite. According to one variety of the present invention, these programs are embodied for use in a structured system of data-logic processors (e.g. knowledge-engineering workstation, computer, process-management computer). According to another variety of the present invention, these programs are embodied for use in a distributed asynchronous system of process-modeling computers.

While the two varieties are architecturally different, the structured system functionally is a substantially hierarchical (graph directed) organization of the same method embodied programs as those of the distributed system. However, because there are substantive differences between command, control, and communications topologies of the structured and asynchronous systems of the present invention, actual method embodied programs conforming to the knowledge-engineering suite may be embodied differently for each system variety.

Furthermore, there are also numerous hybrid, recursive, or quasi-recursive embodiments of the protocol suite of the present invention, which actually constitute interim embodiments between the structured and distributed systems of the present invention. These interim embodiments need not be explicitly described, since substantially they include a mutually compatible aggregation of equivalencies to aspects of the structures system with aspects of the distributed system.

The knowledge-engineering protocol suite of the present invention provides a conceptual organization that is built on the familiar OSI model, and is facilely applied to disparate applications; such as those that differ greatly with respect to “complexity”, “resolution” and “quantification”. The embodied programs of the present invention generally include search-space organizational validation for such disparate applications, and also other higher knowledge-engineering functions. In the protocol-suite, programs provide a synergistic combining of knowledge bases of disparate resolution data-sets, such as by actual or simulated integrating of lower resolution expert-experience based

model-like templates to higher resolution empirical data-capture dense quantitative search-spaces.

The knowledge-engineering protocol suite of the present invention may be applied to disparate applications, such as manufacturing systems, control systems, command control systems, or command control communications systems. Furthermore, the suite may be applied to computational apparatus associated with these applications, and to the task of providing appropriate quantitative modeling and measuring tools.

The present also relates to a **search-space organizational validation method** substantially complying with a knowledge-engineering protocol-suite, the method including the steps of:

A) **organizing** a search-space for a first plurality of correlated empirical data-sets, by **mapping** a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, *at least until there is a predetermined measure of inclusion by the second plurality of nodes and relationships of particulars in the first plurality data-sets,* wherein the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality; and

B) **validating** the search-space from a vantage of a presumption of validity for the first plurality of data-sets, by

I) **simulating** a validity-metric for an n-tuple of directed graph components in the mapped second plurality, or

II) **measuring** if each input to a node of the n-tuple significantly contributes to that nodes output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

In the context of the present protocol-suite, correlated empirical data-sets may are derived from sensors of layer 1, conveyed via a communications conduit facility of layer 2, and stored in a memory media of layer 3. More specifically, correlated empirical data-sets generally include raw input, process, or output data from a specific machine or a specific organism, or from a plurality of specific machines or a plurality of specific organisms, or from a conceptual characterization thereof, or from a simulation of a model relating thereto. According to the domain of problems on which the knowledge-engineering protocol-suite of the present invention operates, illustrative non-limiting examples include:

- A specific machine may be an identified etching machine, or an identified annealing oven in a semiconductor fabrication facility, or an identified locomotive engine, or an identified component or sub-system of a specific machine.
- A specific organism may be an identified individual person, or an identified dairy cow or race horse, or an identified strain of genetically substantially identical bacteria, or an identified organ of any of the aforesaid specific organisms.
- A plurality of specific machines may be a stage in an identified industrial process facility wherein more than one functionally identical specific machines divide portion of a common input into a parallel process and thereafter into a common output. In this context a semiconductor fabrication facility may divide workflow at a specific stage into one of a group of annealing ovens, presumably because annealing is a timely process while other stages of the fabrication are more “instant”. This type of “plurality of specific machines” generally occurs at any stage in an industrial process that would otherwise impose a delay on the entire process, unless such a parallel processing is simultaneously for an excessively economically costly machine.
- A plurality of specific organisms may be a human family, a herd of dairy cows, or even a fermentation vat.

- A conceptual characterization thereof may be a household, a grocery store in a chain of grocery stores, an elementary school or a class therein,.
- A simulation of a model relating thereto may be from an annealing oven modeling, from a line-width etching modeling, from an modeling of public health and epidemic factorizations therein, from a dairy herd management modeling, from a social modeling of parameters in elementary education, etc.

More specifically, interrelated nodes of graph-directed expertise-suggested data-set relationships generally may relate to quantitative or qualitative "axioms" which are either accepted as true in a specific domain of applied knowledge, or are postulated by at least one "expert" according to his long felt suspicions. Diverse situation specific examples of such axioms may include: "Etching line width is primarily dependent on certain specific voltage settings of the etching station" or "An individual cow's milk production is dependent on three specific environmental factors and four specific nutritional factors" or "The fuel efficiency of a locomotive engine seems to degrade when there has been a lot of up-hill acceleration or a lot of down-hill braking". These expertise-suggested data-set relationships are stored on a memory media of layer 3, however these relations are embodied into a topological graph using facilities in layer 4 of the present protocol-suite.

More specifically, *a predetermined measure of inclusion* generally relates to a logical intersection between the first plurality of empirical data-sets (associated with layer 1 of the present protocol) and the second plurality of expertise-suggested relationships (associated with layer 3 of the present protocol). Often there is a disparity of scope between the two pluralities. Either there are extra empirical data-sets that have not been addressed as pertinent to relationships, or there are relationships that do not have supporting data-sets from which to test their validity, or there are both extraneous data-sets and extraneous relationships. It is substantially only in the region of defined

relationships having supporting data-sets that any validation can be attempted. This region must be an interconnected entity at the relationship level. The relationship between disjoint relational sub-sets, even if each is supported by its own respective empirical data-sets, is a problem that may only be addressed in layers 6 or 7 of the present protocol. Hence, a predetermined inclusion specifically relates to a topological sub-graph of relationships that can be validated by virtue of having a sufficient pool of empirical records, which can falsify and test each relationship in the sub-graph according to its respective observed empirical truth. Thus, predetermined in this context relates to a sufficiency for validating according to some statistical metric of certification (e.g. within a first or second standard deviation of average) or some blanket assertion (e.g. this can't happen, or this always happens, or usually this acts in some prescribed fashion).

More specifically, *the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality* relates to a situation where the topological complexity of the expertise suggested relationships is not more complex than the supporting data. One may clarify these cases with three examples. Firstly, the most common acceptable modeling situation describes a small number of inter-related variables that can be tested against a large collection of empirical data. Secondly, a less coherent class of modeling exists when each individual instantiation must be tested against substantially the entire empirical data collection, and this occurs when trying to diagnose and treat an individual patient or when trying to tune an individual racecar. However, to describe a model that captures more relationships than there are n-tuples of empirical data-sets, it is outside the scope of the present invention. For example, to consider validating the truth of a large literary or poetic semantic description (relationship-model) of an individual item (having few associated empirical data-sets) will not allow any cognitive convergence within the present protocol, and accordingly is outside the present scope. In the absence of such a cognitive convergence, the

operations performed in the context of layers 5-7 of the present protocol-suite may prove to be computationally divergent.

More specifically, *a vantage of a presumption of validity* relates to using empirical data in its current form. While many appurtenances may be applied to filter or normalize data, the present invention does not perform these operations. The present invention may be used to characterize an empirical data-set as being statistically distant from other like data-sets. The present invention may used to characterize an individual data instance within a data-set as being statistically distant from other like data-instances. However, these characterizations are of secondary importance in the context of the objects of the present invention. The operational postulate of the present invention is that a model, as composed from individual or collective expertise, may be validated and improved when considered in juxtaposition to empirical data. In the context of the present invention, a data anomaly is an object of study and analysis, not a target for correction. The present invention has an object of finding out what relations characterize this empirical anomaly. It may be that this anomaly is a false representation of the empirical reality. Alternatively, it may be that this anomaly is a statistically rare representative instance of some combination of relationships that might contribute to broadening understanding in the context of a system under study. It is a salient feature of the present invention to disclose and investigate such rare representative instances. Therefore it would be at cross-purposes to the present invention to automatically filter out the very instances that might be most productive to improving knowledge of a system under study.

More specifically, *a validity-metric* relates to a synthetic scale assignment that is derived when a relationship or aggregation of relationships is quantitatively evaluated according to the empirical data. Generally, the metric may reflect a reality that an expertise-suggested relationship is completely supported by the data, or that the relationship only accounts for or correlates

with some measurable part of the data, or that the relationship is not supported by the data, or even that the data supports a relationship contrary to that suggested by an "expert".

More specifically, *n-tuple* relates to a "multiple of n" ("n" being two or more). In the context to the simulating operation, an n-tuple relates to one or more relations between two or more nodes in a directed graph representation for the expertise-suggested data-set relationships.

More specifically, *significantly contributes* relates to another validity metric. Just as for the case of validity metric it is important to know if the empirical data supports the expertise-suggested relationships (between an n-tuple of nodes), it is also important to know this same metric from the vantage of a single node. From the vantage of a single node, one can measure if there is a causal relation between input factors and output results. For example, a node in a process may have temperature designated as a significant input factor to the quality of output products from that node, and this may not necessarily be the case when considered empirically.

More specifically, the method steps of the **search-space organizational validation method** relate to:

A) **organizing** a search-space for a first plurality of correlated empirical data-sets, by **mapping** a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, *at least until there is a predetermined measure of inclusion by the second plurality of nodes and relationships of particulars in the first plurality data-sets,* wherein the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality; and

B) **validating** the search-space from a vantage of a presumption of validity for the first plurality of data-sets, by

I) **simulating** a validity-metric for an n-tuple of directed graph components in the mapped second plurality, or

II) using the validity-metric, **measuring** if each input to a node of the n-tuple significantly contributes to that nodes output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

The present invention further relates to **a program storage device** readable by a logic-machine, tangibly embodying a program of instructions executable by the logic-machine (e.g. a data-logic processor or a process-modeling computer) to perform method steps for validating a search-space organization substantially complying with a knowledge-engineering protocol-suite, said method steps including:

A) **organizing** a search-space for a first plurality of correlated empirical data-sets, by **mapping** a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, at least until the second plurality of nodes and relationships substantially includes a predetermined measure of particulars in the first plurality data-sets, wherein data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality; and

B) **validating** the search-space from a vantage of a presumption of validity for the first plurality of data-sets, by

I) **simulating** a validity-metric for an n-tuple of directed graph components in the mapped second plurality, or

II) using the validity-metric, **measuring** if each input to a node of the n-tuple significantly contributes to that nodes output, wherein

a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

Likewise, the present invention relates to **a process-modeling computer** for use in a distributed asynchronous system of process-modeling computers substantially according to a knowledge-engineering protocol-suite, the process-modeling computer logically having three active-units wherein each active-unit has at least one virtual computer processor associated therewith and wherein the active-units are capable of mutual data-communications interaction, and the process-modeling computer includes:

A) a first active-unit of the three active-units, and said first active-unit is further capable of data-communications interaction with

I) sensors or actuators of an associated process-control machine,

II) at least one other process-modeling computer in the system of process-modeling computers, and

III) at least one data storage device wherein is collectively represented on at least one memory medium

(i) a first plurality of correlated empirical data-sets including at least one data-set of empirical data for the associated process-control machine, and

(ii) a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships

(One) wherein the second plurality includes a directed graph component to or from a representation for the associated process-control machine, and

(Two) wherein the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality;

B) a second active-unit of the three active-units, and said second active-unit is capable of organizing a search-space, for the first plurality of correlated empirical data-sets from the vantage of the associated process-control machine, by mapping, the second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, at least until the second plurality of nodes and relationships substantially includes

I) a predetermined measure of particulars in the at least one data-set of empirical data for the associated process-control machine, and

II) from the relationships, all directed graph components to or from the associated process-control machine; and

C) a third active-unit of the three active-units, and said third active-unit is capable of validating the search-space by

I) simulating a validity-metric for at least one n-tuple of directed graph components in the mapped second plurality, wherein each said n-tuple includes a directed graph component to or from the associated process-control machine, or

II) measuring if each input to a node of the n-tuple significantly contributes to that node's output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

Furthermore, the present invention relates to **a distributed asynchronous system of process-modeling computers** substantially complying with a

knowledge-engineering protocol-suite, the system of process-modeling computers including:

A) at least one process-modeling terminal wherein at least one of the terminals includes a program storage device as described (above);

B) a plurality of process-modeling computers wherein each computer is as described (immediately above);

C) a data-communications interaction conduit providing sufficient transactional data exchange services

I) between the plurality of process-modeling computers;

II) between at least one of the process-modeling terminals and the plurality of process-modeling computers; and

III) between the process-modeling terminals.

In order to further facilitate a coherent appreciation of the broad aspects of present invention, an analogous comparison to the well-known **ISO** (International Standards Organization) **OSI** (Open Systems Interconnection) **reference model** must be noted. This analogous construction is used to organize the many interrelated aspects of the present invention, and to emphasize which outstanding needs of the prior art are beneficially invigorated thereby.

The well-known **ISO** (International Standards Organization) **OSI** (Open Systems Interconnection) **reference model** describes a broad categorization for protocol-suites and elements therein (described the General Background Of The Invention section, and more specifically cited references). This model has proved to be very useful in a broad spectrum of data-communications applications; especially for understanding the structure of large systems; and furthermore for developing operable standards for each part of such structures. Other reference models have been developed (e.g. the US Department of Defense's four-layer reference model). While these other reference models may be conceptually equivalent functionally, they have not proved to be as convenient as the OSI reference model; especially with respect to the interrelated aspects of expressing standards and facilitating understanding.

In a larger context than the ISO reference model, the present invention relates to a knowledge-engineering protocol-suite for facilitating open systems interconnection transactions in a seven–layer reference model. This knowledge-engineering protocol-suite includes: **either** *firstly* a process modeling computer for relating layers 1-3, *secondly* a search-space organizational validation method for relating layers 3-5, and *thirdly* a knowledge-engineering work station for relating layers 5-7; **or** *equivalently* a distributed asynchronous system of process modeling computers for relating layers 1-7.

Therefore, the present invention will be described as including 11 aspects:

- seven mono-layer aspects (each corresponding to a single layer of the suite of the present invention),
- three tri-layer meta-aspects (specifically layers 1-3, 3-5, and 5-7 of the suite of the present invention), and
- one septa-layer mega-aspect (being especially useful for appreciating substantially-decentralized implementations of the suite of the present invention).

These 11 aspects substantially correspond to respective computer programs, systems of computer programs, and computer architectures integrating these systems.

The seven–layer reference model for facilitating open systems interconnection transactions is defined in the context of the present invention as having: a seven layer knowledge-engineering protocol-suite wherein:

LAYER 1 - Layer 1 relates to embodiments of a **physical** layer from which data about physical input, process, or output attributes is collected or targeted. The physical layer may be tied to a physical machine such as a process-controlled machine. The physical layer may be tied to a data input terminal through which input, process, or output data may be collected. The physical layer may be tied to a data output terminal (or a printer) through which input, process, or output transactions may be targeted, reports generated, work-orders authorized, process-control parameters

modified, etc. In all of these examples, the physical layer is tied to an accessible data storage media.

LAYER 2 - Layer 2 relates to embodiments of a **data-link** layer data-communications (e.g. including the ISO OSI model type data-communications per se, inter-net, intra-net, WAN, LAN, DBMS).

LAYER 3 - Layer 3 relates to embodiments of a data-set **network** layer having therein the first plurality data-sets, the second plurality data sets, and other data banks which may yield content that can be manually or automatically transformed into the aforesaid pluralities.

LAYER 4 - Layer 4 relates to a **transport** layer wherein token correspondence (adjacency list) constructions mapping within each plurality and between sets of the pluralities.

LAYER 5 - Layer 5 relates to a **session** layer wherein validation or simulation of the layer 4 mappings may be run on layer 3 data , or as an on the fly control system on layer 1 data.

LAYER 6 - Layer 6 relates to a **presentation** layer where in design of experiments may be articulated for specific sessions.

LAYER 7 - Layer 7 relates to an **application** layer wherein a broader construction of experimental strategy may be articulated such as an n-tuple strategy.

Furthermore, in the context of more preferred scale embodiments of the present invention, the knowledge-engineering protocol-suite pertains to:

- three tri-layer meta-aspects (specifically layers 1-3, 3-5, and 5-7 of the suite of the present invention respectively integrated) that relate to three large embodiments of the present invention, and
- one septa-layer mega-aspect (being especially useful for appreciating substantially-decentralized implementations of the suite of the present invention) that relates to the preferred embodiments of the present invention.

BRIEF DESCRIPTION OF THE FIGURES AND APPENDICES

In order to understand the invention and to see how it may be carried out in practice, embodiments including the preferred embodiment will now be described, by way of non-limiting example only, with reference to the accompanying drawings; in which Figures 1-29 are schematic presentations – specifically:

Figure 1 illustrates systems complying with a knowledge-engineering prorocol-suite;

Figure 2 illustrates apparatus included in the systems of figure 1;

Figure 3 illustrates optional layer 2 protocols for use in the systems of figure 1;

Figure 4 illustrates useful data-ensembles in the context of the systems of figure 1;

Figure 5 illustrates localization of graph-theoretic orderings in the context of the systems of figure 1;

Figure 6 illustrates a program storage device;

Figure 7 illustrates an article of manufacture;

Figure 8 illustrates a process-modeling computer;

Figure 9 illustrates a distributed asynchronous system of process-modeling computers;

Figure 10 illustrates a method of search space organizational validation;

Figures 11-15 illustrate variations of the method of figure 10;

Figures 16-19 illustrate variations of the methods of figures 14-15;

Figures 20-23 illustrate further variations of the method of figure 10;

Figures 24-26 illustrate variations of the method of figure 23;

Figure 27 illustrates another variation option for use with the method of figure 10;

Figure 28 illustrates a variation option for use with the method of figure 13; and

Figure 29 illustrates still another useful variation for use with the method of figure 10.

Illustration 1 portrays a typical schematic knowledge-tree representation example; and

Illustration 2 (A-D) portrays a typical schematic analysis diagram for a conditional SPC example.

Appendix 1 presents, on a CD-ROM, a working prototype of a system embodying aspects of the present invention; and includes an organized collection of source code, documentation thereof, sample menus, and other working appurtenances that have been developed for use therewith; and

Appendix 2 presents, on the same CD-ROM, source code independent descriptive notes and other working papers that have been written in the course of the development of the prototype of appendix 1.

DETAILED DESCRIPTION OF THE FIGURES AND APPENDICES

Figure 1 relates to **a knowledge-engineering protocol-suite** for facilitating open systems interconnection transactions in a multi-layer knowledge-engineering reference model substantially having

Layer 1 – (1/1) a physical layer for interfacing with apparatus (e.g. 2/1);

Layer 2 – (1/2) (1/2a) a data-link layer for facilitating data-communications within any of these Layers 1-7 or between any plurality of these Layers 1-7;

Layer 3 – (1/3) a network layer for maintaining transactional access to data ensembles;

Layer 4 – (1/4) a transport layer for organizing and maintaining token correspondences and adjacency lists wherein are represented network layer relationships between the data sets or between elements in the data sets;

Layer 5 – (1/5) a session layer for validating the transport layer represented relationships and for simulating alternative transport layer relationships;

Layer 6 – (1/6) a presentation layer for designing and executing experimental session layer simulations, evaluations thereof and modifications thereto; and

Layer 7 – (1/7) an application layer for prioritizing n-tuple strategy dynamics of presentation layer transactions;

wherein the knowledge-engineering protocol-suite includes:

either a structured system (1/8) having

at least one process-management computer (1/9) with a program (1/10) for relating Layers 1-3,

at least one computer (1/11) embodying a search-space organizational validation method program (1/12) for relating Layers 3-5, and

at least one knowledge-engineering workstation (1/13) with a program (1/14) for relating Layers 5-7; **or equivalently**

a distributed asynchronous system (1/15) of process-modeling computers (1/16) (1/16a) with programs (1/17) (1/17a) for relating Layers 1-7.

Figure 2 relates to the protocol-suite, as was illustrated in figure 1, wherein the process-management computer or a process-modeling computer includes apparatus (2/1) interfacing with the physical layer, used by the process-management computer or by the distributed asynchronous system of process-modeling computers, and these apparatus are selected from data-communications devices (2/2) or process-control machines (2/3), and the data-communications devices are for input (2/4) or data storage (2/5) or output (2/6), and the process-control machines have sensors (2/7) or program storage (2/8) or actuators (2/9).

Figure 3 relates to the protocol-suite as was illustrated in figure 1 wherein any said program (e.g. (1/10) (1/17) (1/17a)) relating to the data-link layer, used by the process-management computer (e.g. (1/9)) or by the computer (e.g.(1/11)) embodying a search space organizational validation method or by the knowledge-engineering workstation (e.g.

(1/13)) or by the distributed asynchronous system (e.g. (1/15)) of process-modeling computers (e.g. (1/16) (1/16a)), and used for facilitating data-communications within any of the layers 1-7 or between any plurality of the layers 1-7 as required therein, includes at least one data communications protocol (3/1) selected from the list:

ISO OSI model type protocol (3/2),
inter-net type protocol (3/3),
intra-net type protocol (3/4),
Wide Area Network type protocol (3/5),
Local Area Network type protocol (3/6),
Data Base Management System type protocol (3/7),
inter-processor type protocol (3/8),
intra-processor type protocol (3/9).

Figure 4 relates to the protocol-suite as was illustrated in figures 1 and 2 wherein any said program (e.g. (1/10) (1/12) (1/17) (1/17a)) relating to the network layer, used by the process-management computer (e.g. (1/9)) or by the computer (e.g. (1/11)) embodying a search space organizational method or by the distributed asynchronous system (e.g. (1/15)) of process-modeling computers (e.g. (1/16) (1/16a)), and used for maintaining transactional access to data ensembles (4/1), includes in said data ensembles

a first plurality of correlated empirical data-sets (4/2) (4/2a) substantially derived from the process-control machines (e.g. (2/3)) and

a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships (4/3) (4/3a) substantially derived from the data-communications devices (e.g. (2/2)).

Figure 5 relates to the protocol-suite as was illustrated in figure 1 wherein any said program (e.g. (1/14) (1/17) (1/17a)) relating to the application layer, used by the knowledge-engineering workstation (e.g. (1/13)) or by the

distributed asynchronous system (e.g. (1/15)) of process-modeling computers (e.g. (1/16) (1/16a)), and used for (5/1) prioritizing n-tuple strategy dynamics of presentation layer transactions as required therein, includes performing graph-theoretic orderings (5/2) of elements or of sets, and said orderings are performed sequentially, in parallel, concurrently, synchronously, asynchronously, heuristically, or recursively.

Figure 6 relates to a program storage device (6/1) readable by a logic-machine (6/2), tangibly embodying a program (e.g. (1/12)) of instructions executable by the logic-machine to perform method steps for validating a search-space organization substantially complying with a knowledge-engineering protocol-suite, said method steps including:

organizing (6/3) a search-space for a first plurality of correlated empirical data-sets, by **mapping** (6/4) a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, at least until the second plurality of nodes and relationships substantially includes a predetermined measure of particulars in the first plurality data-sets, wherein data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality; and

validating (6/5) the search-space from a vantage of a presumption of validity for the first plurality of data-sets, by

simulating (6/6) a validity-metric for an n-tuple of directed graph components in the mapped second plurality, or

measuring (6/7) if each input to a node of the n-tuple significantly contributes to that nodes output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

Figure 7 relates to an article of manufacture (7/1) including a computer usable medium (7/2) having computer readable program code (7/3)

embodied therein a method for validating a search-space organization and substantially complying with a knowledge-engineering protocol-suite, the computer readable program (e.g. (1/12)) code in said article of manufacture including:

computer readable program code (7/4) for causing a computer to **organize** a search-space for a first plurality of correlated empirical data-sets, by **mapping** a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, at least until the second plurality of nodes and relationships substantially includes a predetermined measure of particulars in the first plurality data-sets, wherein the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality; and

computer readable program code (7/5) for causing the computer to **validate** the search-space from a vantage of a presumption of validity for the first plurality of data-sets, by

simulating a validity-metric for an n-tuple of directed graph components in the mapped second plurality or

measuring if each input to a node of the n-tuple significantly contributes to that nodes output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

Figure 8 relates to a process-modeling computer (1/16) for use in a distributed asynchronous system (e.g. (1/15)) of process-modeling computers substantially according to a knowledge-engineering protocol-suite, the process-modeling computer logically having three active-units (8/1) (8/2) (8/3) wherein each active-unit has at least one virtual computer processor associated therewith (8/1a) (8/2a) (8/3a) and wherein the active-units are capable of mutual data-communications interaction, and the process-modeling computer includes:

a first active-unit (8/1) of the three active-units, and said first active-unit is further capable of data-communications interaction with

sensors (e.g. (2/7)) or actuators (e.g. (2/9)) of an associated process-control machine (e.g. (2/3)),

at least one other process-modeling computer (e.g. (1/16a)) in the system of process-modeling computers, and

at least one data storage device (8/4) wherein is collectively represented on at least one memory medium

a first plurality of correlated empirical data-sets including at least one data-set of empirical data for the associated process-control machine (e.g. (2/3)), and

a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships

wherein the second plurality includes a directed graph component to or from a representation for the associated process-control machine, and

wherein the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality;

a second active-unit (8/2) of the three active-units, and said second active-unit is capable of organizing a search-space, for the first plurality of correlated empirical data-sets from the vantage of the associated process-control machine, by mapping, the second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, at least until the second plurality of nodes and relationships substantially includes

a predetermined measure of particulars in the at least one data-set of empirical data for the associated process-control machine, and

from the relationships, all directed graph components to or from the associated process-control machine; and

a third active-unit (8/3) of the three active-units, and said third active-unit is capable of validating the search-space by

simulating a validity-metric for at least one n-tuple of directed graph components in the mapped second plurality, wherein each said n-tuple includes a directed graph component to or from the associated process-control machine, or

measuring if each input to a node of the n-tuple significantly contributes to that node's output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

Figure 9 relates to a distributed asynchronous system (1/15) of process-modeling computers substantially complying with a knowledge-engineering protocol-suite, the system of process-modeling computers including:

at least one process-modeling terminal (9/1) (9/1a) wherein at least one of the terminals includes a program storage device (6/1) as was illustrated in figure 6;

a plurality of process-modeling computers (1/16) (1/16a) wherein each computer is as was illustrated in greater detail figure 8;

a data-communications interaction conduit (9/2) providing sufficient transactional data exchange services

between the plurality of process-modeling computers;

between at least one of the process-modeling terminals and the plurality of process-modeling computers; and

between the process-modeling terminals.

Figure 10 relates to a search-space organizational validation method (10/1) substantially complying with a knowledge-engineering protocol-suite, the method including the steps of:

organizing (10/2) a search-space for a first plurality of correlated empirical data-sets, by **mapping** (10/3) a second plurality of

interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, *at least until there is a predetermined measure of inclusion by the second plurality of nodes and relationships of particulars in the first plurality data-sets,* wherein the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality; and

validating (10/4) the search-space from a vantage of a presumption of validity for the first plurality of data-sets, by

simulating (10/5) a validity-metric for an n-tuple of directed graph components in the mapped second plurality, or

measuring (10/6) if each input to a node of the n-tuple significantly contributes to that nodes output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

Figure 11 relates to the method as was illustrated in figure 10 wherein **mapping** (10/3) includes **defining** (11/1) substantially every node in the second plurality to have at least one graph-directed input and at least one graph-directed output.

Figure 12 relates to the method as was illustrated in figure 10 wherein **mapping** (10/3) includes **defining** (12/1) substantially every node in the second plurality to have only one graph-directed output.

Figure 13 relates to the method as was illustrated in figure 10 wherein **mapping** (10/3) includes **standardizing** (13/1) a format representation for nodes or relationships in the second plurality.

Figure 14 relates to the method as was illustrated in figure 10 wherein **mapping** (10/3) includes **representing** (14/1) graph-directed data-set relationships using expertise-suggested initial weightings.

Figure 15 relates to the method as was illustrated in figure 10 wherein **mapping** (10/3) includes **representing** (15/1) graph-directed data-set relationships using initial weightings based on statistical process-control generated distribution functions.

Figure 16 relates to the method as was illustrated in figure 14 wherein **validating** (10/4) includes, for at least one weighted directed graph component in the directed graph of second plurality components, **improving** (16/1) the weighted component using a validity-metric proportional directed graph component weighting.

Figure 17 relates to the method as was illustrated in figure 15 wherein **validating** (10/4) includes, for at least one weighted directed graph component in the directed graph of second plurality components, **improving** (17/1) the weighted component using a validity-metric proportional directed graph component weighting..

Figure 18 relates to the method as was illustrated in figure 16 wherein **validating** (10/4) includes **generating** (18/1) a conditional statistical process-control distribution functions and **convoluting** (18/2) the conditional distribution functions with the present weightings.

Figure 19 relates to the method as was illustrated in figure 17 wherein **validating** (10/4) includes **generating** (19/1) a conditional statistical process-control distribution functions and **convoluting** (19/2) the conditional distribution functions with the present weightings.

Figure 20 relates to the method as was illustrated in figure 10 wherein **validating** (10/4) includes, for at least one directed graph component in

the directed graph of second plurality components, **assigning** (20/1) a validity-metric proportional directed graph component weighting.

Figure 21 relates to the method as was illustrated in figure 10 wherein **validating** (10/4) includes, for at least one validity-metric above a threshold value, **adding** (21/1) a virtual directed graph component to the second plurality.

Figure 22 relates to the method as was illustrated in figure 10 wherein **validating** (10/4) includes, for at least one validity-metric below a threshold value, **deleting** (22/1) a directed graph component from the second plurality.

Figure 23 relates to the method as was illustrated in figure 10 wherein **mapping** (10/3) includes **updating** (23/1) the first plurality of correlated empirical data-sets.

Figure 24 relates to the method as was illustrated in figure 23 wherein **updating** (23/1) includes **modifying** (24/1) at least one real-time empirical data-set.

Figure 25 relates to the method as was illustrated in figure 23 wherein **mapping** (10/3) includes **activating** (25/1) an alarm when an updated empirical value is outside of a threshold range.

Figure 26 relates to the method as was illustrated in figure 23 wherein **validating** (10/4) includes **generating** (26/1) a report having recorded therein an updated empirical value that is outside of a threshold range.

Figure 27 relates to the method as was illustrated in figure 10 wherein **mapping** (10/3) includes **accumulating** (27/1) empirical data using a data mining engine.

Figure 28 relates to the method as was illustrated in figure 13 wherein **standardizing** (13/1) a format representation for nodes or relationships in the second plurality includes either **providing** (28/1) for substantially each node in the second plurality: at least one input token; a process token; and at least one output token; or **providing** (28/2) for substantially each relationship in the second plurality: a first process token, a linkage token; and a next process token.

Figure 29 relates to the method as was illustrated in figures 10 wherein **mapping** (10/3) includes **defining** (29/1) a correspondence in the search-space, between the second plurality of interrelated nodes and a process model representation, by performing the steps of:

standardizing (29/2) a format representation for nodes or relationships in the second plurality by

either **providing** (29/3) for substantially each node in the second plurality:

at least one input token;

a process token; and

at least one output token;

or **providing** (29/4) for substantially each relationship in the second plurality:

a first process token,

a linkage token; and

a next process token;

standardizing (29/5) a format representation for nodes or relationships in the process model by

either **providing** (29/6) for substantially each node in the process model:

at least one input token;

a process token; and

at least one output token;

or **providing** (29/7) for substantially each relationship in the process model:

a first process token,

a linkage token; and

a next process token; and

providing (29/8) correspondence rules (29/8a) between **tokens (29/8b) of the second plurality and** tokens (29/8c) of the process model.

Illustration 1 portrays a typical schematic knowledge-tree representation example; and

Illustration 2 (A-D) portrays a typical schematic analysis diagram for a conditional SPC example.

Appendix 1 presents, on a CD-ROM, a working prototype of a system embodying aspects of the present invention; and includes an organized collection of source code, documentation thereof, sample menus, and other working appurtenances that have been developed for use therewith; and

Appendix 2 presents, on the same CD-ROM, source code independent descriptive notes and other working papers that have been written in the course of the development of the prototype of appendix 1.

More specifically, the index of the CD_ROM – relating to 104 File totaling 60,124,848 bytes - is:

- Directory of P:\Patents\prototype and documents CD

- 10/31/99 01:58p <DIR> .
- 10/31/99 01:58p <DIR> ..
- 10/31/99 02:00p 6,361 dirs.txt

- 10/31/99 02:29p <DIR> documents
- 10/31/99 12:34p <DIR> software
- 5 File(s) 6,361 bytes

- Directory of P:\Patents\prototype and documents CD\documents

- 10/31/99 02:29p <DIR> .
- 10/31/99 02:29p <DIR> ..
- 05/13/99 05:50p 151,040 databaseReport.doc
- 10/06/99 08:40p 418,304 eden.doc
- 10/25/99 11:14a 201,728 EdenSummary.doc
- 05/26/99 03:21p 115,712 Gui.doc
- 10/31/99 01:51p <DIR> patents
- 10/31/99 02:40p 572,928 poem4_1.ppt
- 05/27/99 04:59p 964,096 ScientistTechnicalReport.doc
- 10/31/99 01:31p 21,504 SoftwarePatent.doc
- 10 File(s) 2,445,312 bytes

- Directory of P:\Patents\prototype and documents CD\documents\patents

- 10/31/99 01:51p <DIR> .
- 10/31/99 01:51p <DIR> ..
- 06/02/99 09:47a 83,456 Adam and Eve.doc
- 10/27/99 06:46p 20,992 Assignment.doc
- 05/27/99 03:11p 83,456 Empirical Controller Patent-2 OLD.doc
- 05/26/99 02:11p 81,408 Empirical Controller Patent-21 OLD.doc
- 05/28/99 01:07p 150,016 Empirical Controller Patent-3 OLD.doc

- 05/31/99 01:58p 156,672 Empirical Controller Patent-4 OLD.doc
- 06/02/99 09:43a 156,672 Empirical Controller.doc
- 05/28/99 03:18p 25,088 Hartman-Patent Claims.doc
- 05/23/99 07:43a 155,648 IDM System Application File OLD.doc
- 06/23/99 08:18a 35,328 InSyst versus KnowledgeScape.doc
- 06/07/99 04:05p 27,648 INTEL CITR.doc
- 07/12/99 04:36p 25,600 Knowledge Tree description and claims.doc
- 06/02/99 04:43p 30,208 Patent List.doc
- 06/02/99 10:02a 29,184 Patent List_2.doc
- 04/12/99 04:31p 27,136 Patent Summaary- InSyst 4 Primary Patents.doc
- 12/28/98 02:53a 198,144 PATENT_1.doc
- 04/16/99 02:29p 198,144 PATENT_1_1.DOC
- 06/02/99 01:43p 40,960 POEM SPC.doc
- 06/02/99 04:32p 41,472 POEM.doc
- 06/16/99 02:25p 26,624 references.doc
- 07/29/99 02:24p 20,480 Sandy Colb Provisional Patent Application_1.doc
- 08/03/99 03:40p 20,992 Sandy Colb Uniqueness of Knowledge Tree.doc
- 07/14/99 10:16a 110,592 Sandy Meeting Garden Tree.doc
- 25 File(s) 1,745,920 bytes

- Directory of P:\Patents\prototype and documents CD\software

- 10/31/99 12:34p <DIR> .
- 10/31/99 12:34p <DIR> ..

- 10/31/99 01:59p <DIR> exe
- 10/07/99 09:27a 21,504 SoftwarePatent.doc
- 4 File(s) 21,504 bytes

- Directory of P:\Patents\prototype and documents CD\software\exe

- 10/31/99 01:59p <DIR> .
- 10/31/99 01:59p <DIR> ..
- 02/09/98 02:00a 29,952 borlndmm.dll
- 06/22/99 10:26a 15,993 ChartDLL.cpp
- 06/22/99 11:35a 14,336 ChartDLL.dll
- 06/22/99 09:38a 2,465 ChartItExport.h
- 08/25/99 11:53a 477,184 clean.DB
- 02/09/98 02:00a 996,872 cp3240mt.dll
- 01/27/99 03:00a 908,800 cp3245mt.dll
- 08/25/99 12:08p 66,048 dddw.pbd
- 08/25/99 12:05p 96,768 dddw.pbl
- 08/25/99 12:08p 116,224 func_use.pbd
- 08/25/99 12:06p 180,224 func_use.pbl
- 10/24/99 02:38p 19,840,000 insyst.db
- 08/25/99 12:08p 227,840 insyst.exe
- 08/25/99 04:56p 634 insyst.ini
- 08/25/99 12:08p 604,160 insyst.pbd
- 08/25/99 04:38p 1,096,192 insyst.pbl
- 08/25/99 11:28a 386 insyst.pbr
- 08/25/99 12:06p 19,968 ins_general.pbl
- 08/25/99 12:08p 204,288 ins_pfe.pbd
- 08/25/99 12:06p 386,560 ins_pfe.pbl
- 08/25/99 11:37a 52,736 ins_srv.pbl
- 08/25/99 04:35p 113,664 ipc.pbd

- 08/25/99 04:50p 218,624 ipc.pbl
- 08/19/99 09:46a 7,958 IpcTeeDLL.cpp
- 08/19/99 03:01p 1,683 IpcTeeDLL.h
- 08/24/99 04:37p 32,256 IpcTeeProj.dll
- 08/25/99 12:08p 142,848 mapper.pbd
- 08/25/99 12:06p 293,376 mapper.pbl
- 03/02/99 05:01a 462,848 olch3d32.dll
- 07/01/98 03:21p 1,608,704 pbdwe60.dll
- 06/10/98 09:09a 288,768 pbodb60.dll
- 06/09/98 11:08p 62,529 pbodb60.ini
- 06/24/98 01:09p 3,408,384 pbvm60.dll
- 08/25/99 12:08p 1,282,048 pfcapsrv.pbd
- 08/25/99 12:08p 3,225,088 pfcapsrv.pbl
- 08/25/99 12:08p 883,200 pfcdwsrv.pbd
- 08/25/99 12:08p 2,341,888 pfcdwsrv.pbl
- 08/25/99 12:08p 995,840 pfcmain.pbd
- 08/25/99 12:07p 2,557,440 pfcmain.pbl
- 08/25/99 12:08p 607,744 pfcutil.pbd
- 08/25/99 12:08p 1,186,304 pfcutil.pbl
- 08/25/99 12:08p 311,808 pfcwnsrv.pbd
- 08/25/99 12:08p 679,936 pfcwnsrv.pbl
- 08/25/99 12:08p 187,904 pfeapsrv.pbd
- 08/25/99 12:07p 235,008 pfeapsrv.pbl
- 08/25/99 12:08p 108,032 pfedwsrv.pbd
- 08/25/99 12:07p 132,096 pfedwsrv.pbl
- 08/25/99 12:08p 190,976 pfemain.pbd
- 08/25/99 12:07p 246,272 pfemain.pbl
- 08/25/99 12:08p 198,144 pfeutil.pbd
- 08/25/99 12:08p 232,960 pfeutil.pbl
- 08/25/99 12:08p 143,872 pfewnsrv.pbd

- 08/25/99 12:08p 175,104 pfewnsrv.pbl
- 10/22/98 03:01a 382,456 tee40.bpl
- 02/09/98 02:00a 5,032,175 vcl35.bpl
- 01/07/99 03:02a 1,888,232 vcl40.bpl
- 08/25/99 12:08p 237,056 wizard_insyst.pbd
- 08/25/99 12:06p 464,896 wizard_insyst.pbl
- 60 File(s) 55,905,751 bytes

- Total Files Listed:
- 104 File(s) 60,124,848 bytes

DETAILED DESCRIPTION OF THE INVENTION

Simply stated, the present invention relates to improving the quality of process control by using expert knowledge which facilitates constructing a topological process graph (often times a directed graph) from the descriptions of at least one expert, or even from a composite collection of interviewing many involved workers (e.g. in situations where not even one expert study has ever been conducted). At this juncture there is a model of a system or process, not unlike models that are constructed in other modeling type systems (described above). **Illustration 1** portrays a typical schematic knowledge-tree representation example of all or part of such a model. Other sample representations may be constructed automatically by running the prototype (of appendix 1) on a sample data base (also in appendix 1) or on another data base of equivalent form.

Since this model is independent of the level of detail that it describes and since this model may capture multiple descriptions that may even contradict each other, the present invention allows the composite model construction to be used in a novel way. Initially, it is important to validate the composite model. Testing each link in the composite model against actual empirical data accomplishes this validation. In the even that a statistically inadequate quanta of empirical data is

available, then the model may be tested against simulation data which was seeded by the empirical data; or in the worst case, by theoretical suppositions.

Validation of each link may be expressed quantitatively. For example, a correlation represented by a link between two nodes may be suppurated by all available data (100% validated), by some lesser plurality of the data, not at all, or even in opposition to the actual empirical data. At this stage, quantative validation may be used to prune out links whose evaluation is below an acceptable threshold. Likewise, the same validation that has been used to test expert suggested relationships may be applied to evaluated new suppositional relationships; and even to substitute suppositional relations of greater validity for expertise suggested relationships of lesser validity.

A variety of strategies may be applied to the task of postulating suppositional relationships. Foremost among these strategies is the application of a new SPC strategy for robust-like optimizations; an example of which is present in **Illustration 2 (A-D);** which portrays a typical schematic analysis diagram for a conditional SPC example. Other sample optimizations may be constructed automatically by running the prototype (of appendix 1) on a sample database (also in appendix 1) or on another database of equivalent form.

The attached prototype, on CD-ROM includes:

For an Environment:

- This system is developed under MS – NT 4.0 operating system.
- The database is SYBASE SQL ANYWARE version 5.5 in stand-alone version.
- The connection to database is via ODBC.
- The source code is written in PowerBuilder version 6.5.
- The C++ code is written in Borland C++ builder version 4.0 with additional dlls of OlectraChart 6.0 charting software.

- A Poem executable file:
- insyst.exe

- Power Builder Libraries developed by InSyst:

- dddw.pbl
- func_use.pbl using system model
- ins_general.pbl general routines
- ins_pfe.pbl
- insyst.pbl
- ipc.pbl calculating and drawing IPC charts
- mapper.pbl creating system model
- wizard_insyst.pbl wizards used in Poem
- dddw.pbd
- func_use.pbd using system model
- ins_pfe.pbd
- insyst.pbd
- ipc.pbd calculating and drawing IPC charts
- mapper.pbd creating system model
- wizard_insyst.pbd wizards used in Poem

- Power Builder Libraries used by InSyst:
- pfcapsrv.pbl
- pfcdwsrv.pbl
- pfcmain.pbl
- pfcutil.pbl
- pfcwnsrv.pbl
- pfeapsrv.pbd
- pfedwsrv.pbd
- pfemain.pbd
- pfeutil.pbld
- pfewnsrv.pbd

- C++ code developed by InSyst:
- ChartDLL.cpp displays 3D response curve
- IpcTeeDLL.cpp displays APC and SPC charts

- DLLs developed by InSyst:
- ChartDLL.dll displays 3D-response curve
- IpcTeeProj.dll displays APC and SPC charts

- DATABASE file:
- Insyst.db

- Other DLLs required:
- borlndmm.dll for C++ code compiled by Borland C++ Builder
- cp3240mt.dll
- cp3245mt.dll
- pbodb60.dll for code written in Power Builder
- pbdwe60.dll for code written in Power Builder
- pbvm60.dll for code written in Power Builder
- olch3d32.dll charting software of OlectraChart

- Additional files required:
- insyst.exe
- vcl35.bpl
- vcl40.bpl
- tee40.bpl

CLAIMS

1. A **knowledge-engineering protocol-suite** for facilitating open systems interconnection transactions in a multi-layer knowledge-engineering reference model substantially having

Layer 1 - a physical layer for interfacing with apparatus;

Layer 2 - a data-link layer for facilitating data-communications within any of these Layers 1-7 or between any plurality of these Layers 1-7;

Layer 3 - a network layer for maintaining transactional access to data ensembles;

Layer 4 - a transport layer for organizing and maintaining token correspondences and adjacency lists wherein are represented network layer relationships between the data sets or between elements in the data sets;

Layer 5 - a session layer for validating the transport layer represented relationships and for simulating alternative transport layer relationships;

Layer 6 - a presentation layer for designing and executing experimental session layer simulations, evaluations thereof and modifications thereto; and

Layer 7 - an application layer for prioritizing n-tuple strategy dynamics of presentation layer transactions;

wherein the knowledge-engineering protocol-suite includes:

A) either a structured system having

I) at least one process-management computer with a program for relating Layers 1-3,

II) at least one computer embodying a search-space organizational validation method program for relating Layers 3-5, and

III) at least one knowledge-engineering workstation with a program for relating Layers 5-7;

B) or equivalently a distributed asynchronous system of process-modeling computers with programs for relating Layers 1-7.

2. The protocol-suite according to claim 1 wherein the process-management computer or a process-modeling computer includes apparatus interfacing with the physical layer, used by the process-management computer or by the distributed asynchronous system of process-modeling computers, and these apparatus are selected from data-communications devices or process-control machines, and the data-communications devices are for input or data storage or output, and the process-control machines have sensors or program storage or actuators.

3. The protocol-suite according to claim 1 wherein any said program relating to the data-link layer, used by the process-management computer or by the computer embodying a search space organizational validation method or by the knowledge-engineering workstation or by the distributed asynchronous system of process-modeling computers, and used for facilitating data-communications within any of the layers 1-7 or between any plurality of the layers 1-7 as required therein, includes at least one data communications protocol selected from the list:

A) ISO OSI model type protocol,
B) inter-net type protocol,
C) intra-net type protocol,
D) Wide Area Network type protocol,
E) Local Area Network type protocol,
F) Data Base Management System type protocol,
G) inter-processor type protocol,
H) intra-processor type protocol.

4. The protocol-suite according to claims 1 and 2 wherein any said program relating to the network layer, used by the process-management computer or by the computer embodying a search space organizational method or by the distributed asynchronous system of process-modeling

computers, and used for maintaining transactional access to data ensembles, includes in said data ensembles

A) a first plurality of correlated empirical data-sets substantially derived from the process-control machines and

B) a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships substantially derived from the data-communications devices.

5. The protocol-suite according to claim 1 wherein any said program relating to the application layer, used by the knowledge-engineering workstation or by the distributed asynchronous system of process-modeling computers, and used for prioritizing n-tuple strategy dynamics of presentation layer transactions as required therein, includes performing graph-theoretic orderings of elements or of sets, and said orderings are performed sequentially, in parallel, concurrently, synchronously, asynchronously, heuristically, or recursively.

6. A program storage device readable by a logic-machine, tangibly embodying a program of instructions executable by the logic-machine to perform method steps for validating a search-space organization substantially complying with a knowledge-engineering protocol-suite, said method steps including:

A) **organizing** a search-space for a first plurality of correlated empirical data-sets, by **mapping** a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, at least until the second plurality of nodes and relationships substantially includes a predetermined measure of particulars in the first plurality data-sets, wherein data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality; and

B) **validating** the search-space from a vantage of a presumption of validity for the first plurality of data-sets, by

I) **simulating** a validity-metric for an n-tuple of directed graph components in the mapped second plurality, or

II) **measuring** if each input to a node of the n-tuple significantly contributes to that nodes output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

7. An article of manufacture including a computer usable medium having computer readable program code embodied therein for validating a search-space organization and substantially complying with a knowledge-engineering protocol-suite, the computer readable program code in said article of manufacture including:

A) computer readable program code for causing a computer to **organize** a search-space for a first plurality of correlated empirical data-sets, by **mapping** a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, at least until the second plurality of nodes and relationships substantially includes a predetermined measure of particulars in the first plurality data-sets, wherein the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality; and

B) computer readable program code for causing the computer to **validate** the search-space from a vantage of a presumption of validity for the first plurality of data-sets, by

I) **simulating** a validity-metric for an n-tuple of directed graph components in the mapped second plurality or

II) **measuring** if each input to a node of the n-tuple significantly contributes to that nodes output, wherein a predetermined

convolution of these measurings constitutes a validity-metric for the n-tuple.

8. A process-modeling computer for use in a distributed asynchronous system of process-modeling computers substantially according to a knowledge-engineering protocol-suite, the process-modeling computer logically having three active-units wherein each active-unit has at least one virtual computer processor associated therewith and wherein the active-units are capable of mutual data-communications interaction, and the process-modeling computer includes:

A) a first active-unit of the three active-units, and said first active-unit is further capable of data-communications interaction with

I) sensors or actuators of an associated process-control machine,

II) at least one other process-modeling computer in the system of process-modeling computers, and

III) at least one data storage device wherein is collectively represented on at least one memory medium

(i) a first plurality of correlated empirical data-sets including at least one data-set of empirical data for the associated process-control machine, and

(ii) a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships

(One) wherein the second plurality includes a directed graph component to or from a representation for the associated process-control machine, and

(Two) wherein the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality;

B) a second active-unit of the three active-units, and said second active-unit is capable of organizing a search-space, for the first plurality of correlated empirical data-sets from the vantage of the associated process-control machine, by mapping, the second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, at least until the second plurality of nodes and relationships substantially includes

I) a predetermined measure of particulars in the at least one data-set of empirical data for the associated process-control machine, and

II) from the relationships, all directed graph components to or from the associated process-control machine; and

C) a third active-unit of the three active-units, and said third active-unit is capable of validating the search-space by

I) simulating a validity-metric for at least one n-tuple of directed graph components in the mapped second plurality, wherein each said n-tuple includes a directed graph component to or from the associated process-control machine, or

II) measuring if each input to a node of the n-tuple significantly contributes to that node's output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

9. A distributed asynchronous system of process-modeling computers substantially complying with a knowledge-engineering protocol-suite, the system of process-modeling computers including:

A) at least one process-modeling terminal wherein at least one of the terminals includes a program storage device according to claim 6;

B) a plurality of process-modeling computers wherein each computer is according to claim 8;

C) a data-communications interaction conduit providing sufficient transactional data exchange services

I) between the plurality of process-modeling computers;

II) between at least one of the process-modeling terminals and the plurality of process-modeling computers; and

III) between the process-modeling terminals.

10. A search-space organizational validation method substantially complying with a knowledge-engineering protocol-suite, the method including the steps of:

A) **organizing** a search-space for a first plurality of correlated empirical data-sets, by **mapping** a second plurality of interrelated nodes of graph-directed expertise-suggested data-set relationships onto the first plurality of correlated empirical data-sets, *at least until there is a predetermined measure of inclusion by the second plurality of nodes and relationships of particulars in the first plurality data-sets,* wherein the data-set resolution of particulars in the first plurality is greater than or equal to that of particulars in the second plurality; and

B) **validating** the search-space from a vantage of a presumption of validity for the first plurality of data-sets, by

I) **simulating** a validity-metric for an n-tuple of directed graph components in the mapped second plurality, or

II) **measuring** if each input to a node of the n-tuple significantly contributes to that nodes output, wherein a predetermined convolution of these measurings constitutes a validity-metric for the n-tuple.

11. The method according to claim 10 wherein **mapping** includes **defining** substantially every node in the second plurality to have at least one graph-directed input and at least one graph-directed output.

12. The method according to claim 10 wherein **mapping** includes **defining** substantially every node in the second plurality to have only one graph-directed output.

13. The method according to claim 10 wherein **mapping** includes **standardizing** a format representation for nodes or relationships in the second plurality.

14. The method according to claim 10 wherein **mapping** includes **representing** graph-directed data-set relationships using expertise-suggested initial weightings.

15. The method according to claim 10 wherein **mapping** includes **representing** graph-directed data-set relationships using initial weightings based on statistical process-control generated distribution functions.

16. The method according to claim 14 wherein **validating** includes, for at least one weighted directed graph component in the directed graph of second plurality components, **improving** the weighted component using a validity-metric proportional directed graph component weighting.

17. The method according to claim 15 wherein **validating** includes, for at least one weighted directed graph component in the directed graph of second plurality components, **improving** the weighted component using a validity-metric proportional directed graph component weighting.

18. The method according to claim 16 wherein **validating** includes **generating** a conditional statistical process-control distribution functions and **convoluting** the conditional distribution functions with the present weightings.

19. The method according to claim 17 wherein **validating** includes **generating** a conditional statistical process-control distribution functions and **convoluting** the conditional distribution functions with the present weightings.

20. The method according to claim 10 wherein **validating** includes, for at least one directed graph component in the directed graph of second plurality components, **assigning** a validity-metric proportional directed graph component weighting.

21. The method according to claim 10 wherein **validating** includes, for at least one validity-metric above a threshold value, **adding** a virtual directed graph component to the second plurality.

22. The method according to claim 10 wherein **validating** includes, for at least one validity-metric below a threshold value, **deleting** a directed graph component from the second plurality.

23. The method according to claim 10 wherein **mapping** includes **updating** the first plurality of correlated empirical data-sets.

24. The method according to claim 23 wherein **updating** includes **modifying** at least one real-time empirical data-set.

25. The method according to claim 23 wherein **mapping** includes **activating** an alarm when an updated empirical value is outside of a threshold range.

26. The method according to claim 23 wherein **validating** includes **generating** a report having recorded therein an updated empirical value that is outside of a threshold range.

27. The method according to claim 10 wherein **mapping** includes **accumulating** empirical data using a data mining engine.

28. The method according to claim 13 wherein **standardizing** a format representation for nodes or relationships in the second plurality includes either **providing** for substantially each node in the second plurality: at least one input token; a process token; and at least one output token; or **providing** for substantially each relationship in the second plurality: a first process token, a linkage token; and a next process token.

29. The method according to claims 10 wherein **mapping** includes **defining** a correspondence in the search-space, between the second plurality of interrelated nodes and a process model representation, by performing the steps of:

A) **standardizing** a format representation for nodes or relationships in the second plurality by

I) either **providing** for substantially each node in the second plurality:

(a) at least one input token;

(b) a process token; and

(c) at least one output token;

II) or **providing** for substantially each relationship in the second plurality:

(a) a first process token,

(b) a linkage token; and

(c) a next process token;

B) **standardizing** a format representation for nodes or relationships in the process model by

I) either **providing** for substantially each node in the process model:

(a) at least one input token;

(b) a process token; and

(c) at least one output token;

II) or **providing** for substantially each relationship in the process model:

(a) a first process token,

(b) a linkage token; and

(c) a next process token; and

C) **providing** correspondence rules between

I) tokens of the second plurality and

II) tokens of the process model.

30. **A knowledge-engineering protocol-suite** substantially as hereinbefore described and illustrated.

For the Applicant, חיים

By: **Chaim Scheff**

Patent Attorney

InSyst Ltd. (Israel)

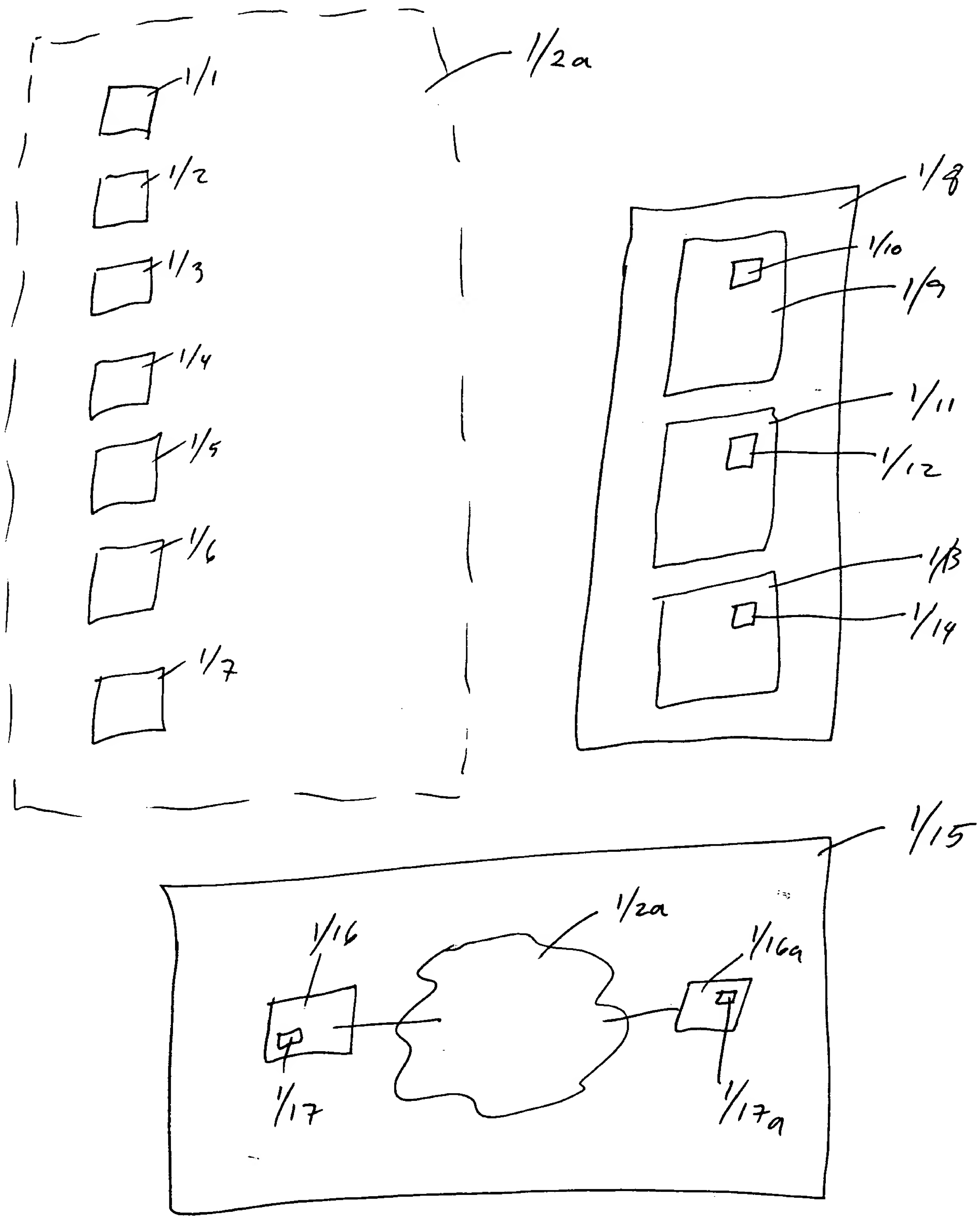
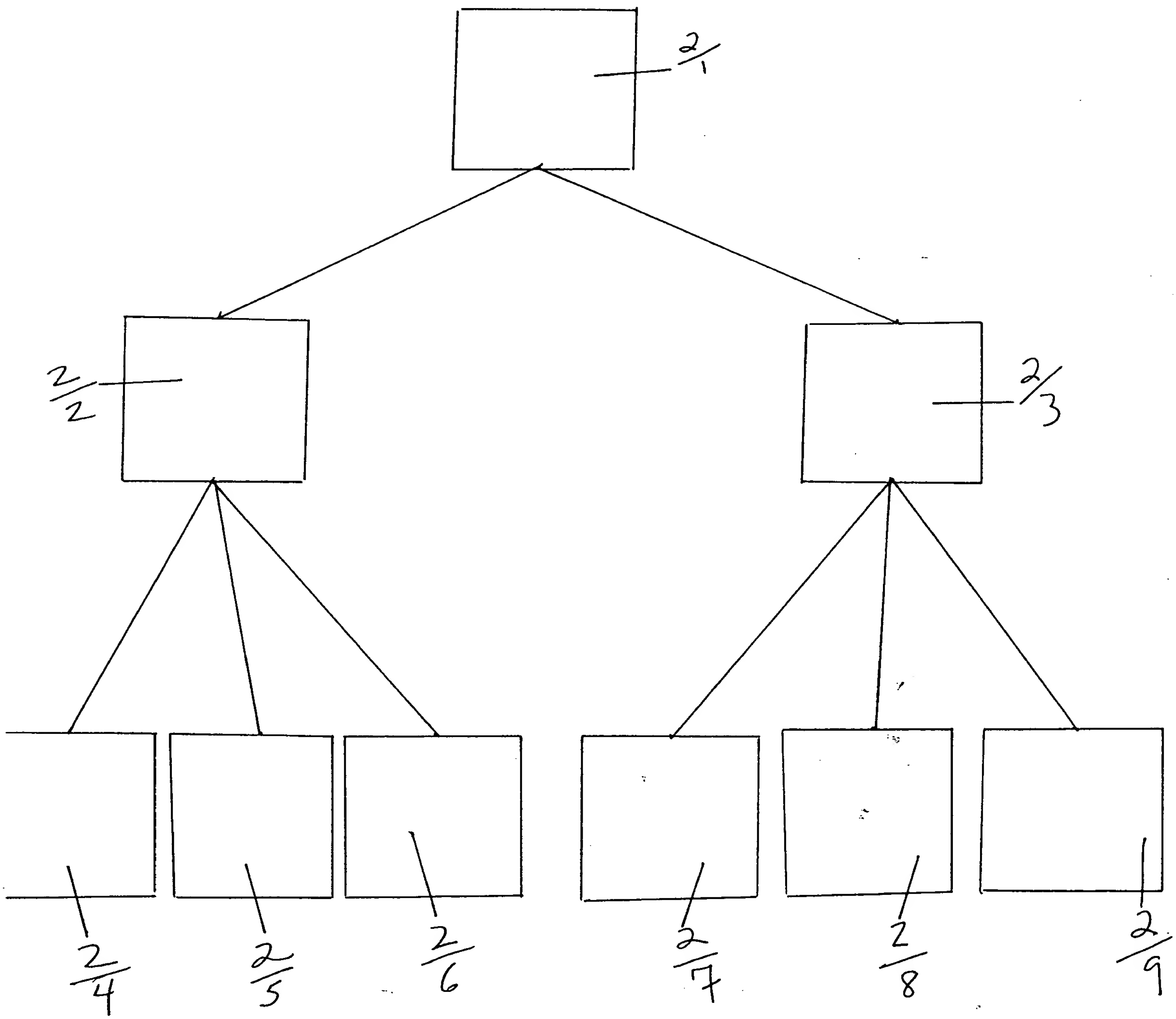

1/2a
1/1
1/2
1/3
1/4
1/5
1/6
1/7
1/8
1/10
1/9
1/11
1/12
1/13
1/14
1/15
1/2a
1/16
1/16a
1/17
1/17a

FIG 1


2/1
2/2
2/3
2/4
2/5
2/6
2/7
2/8
2/9

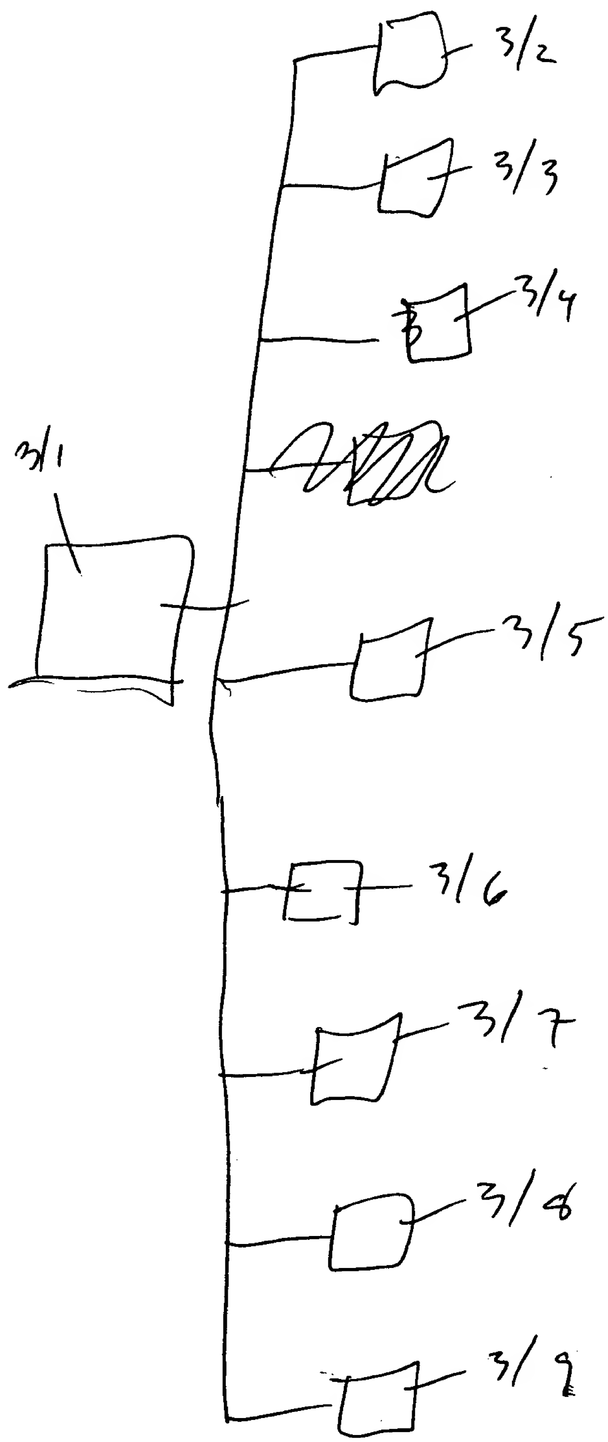


FIG 3

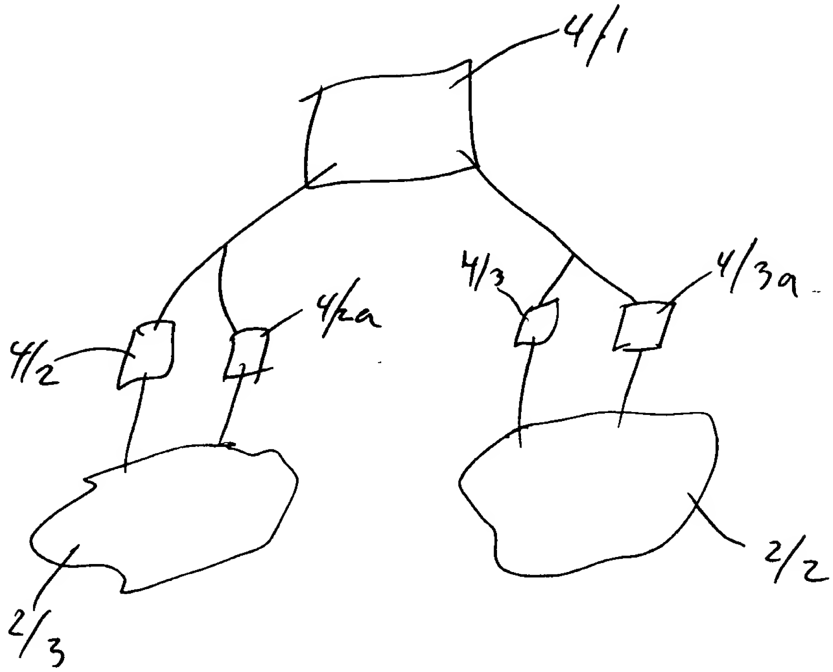


FIG 4

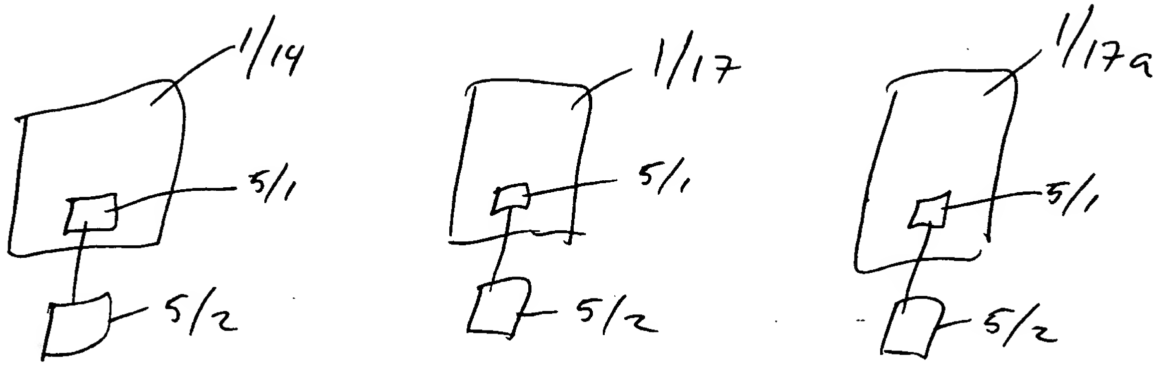


FIG 5

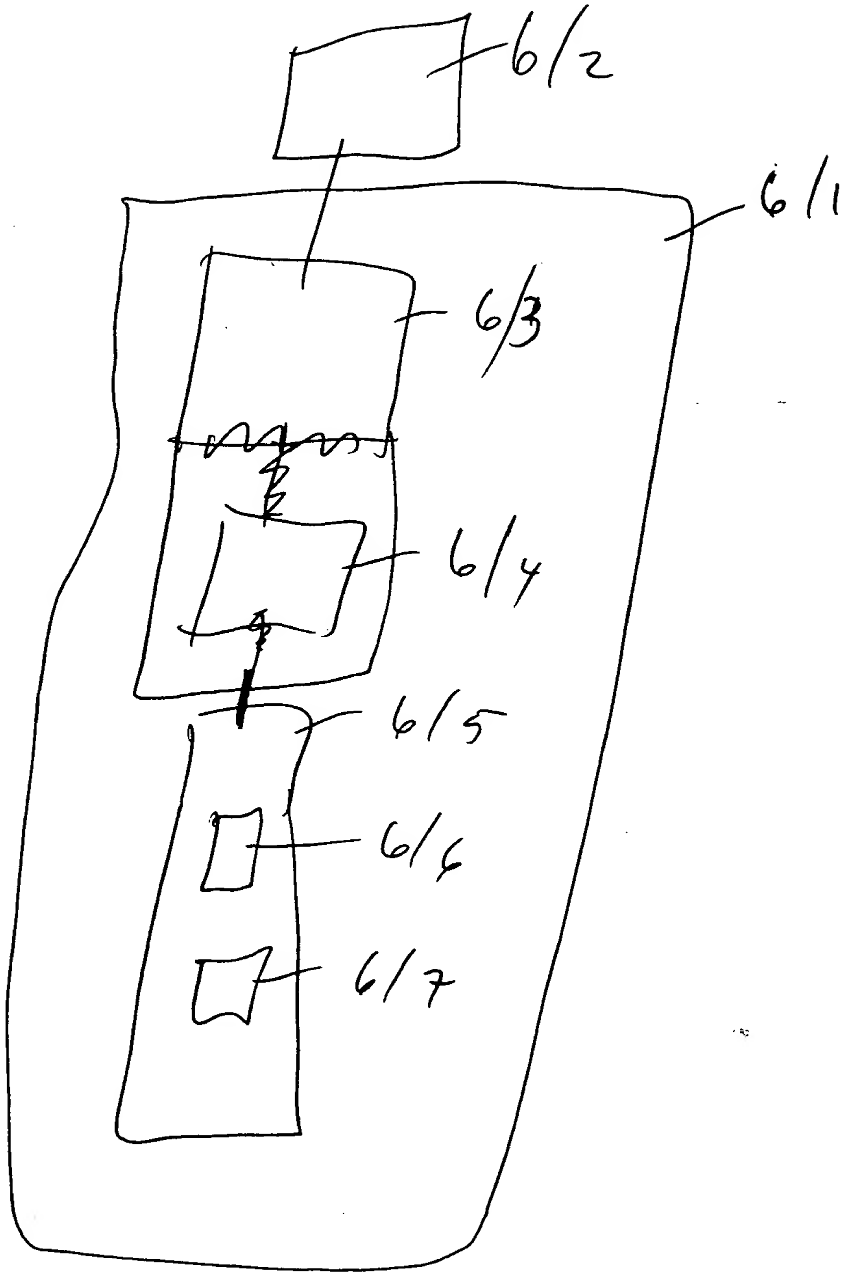


FIG 6

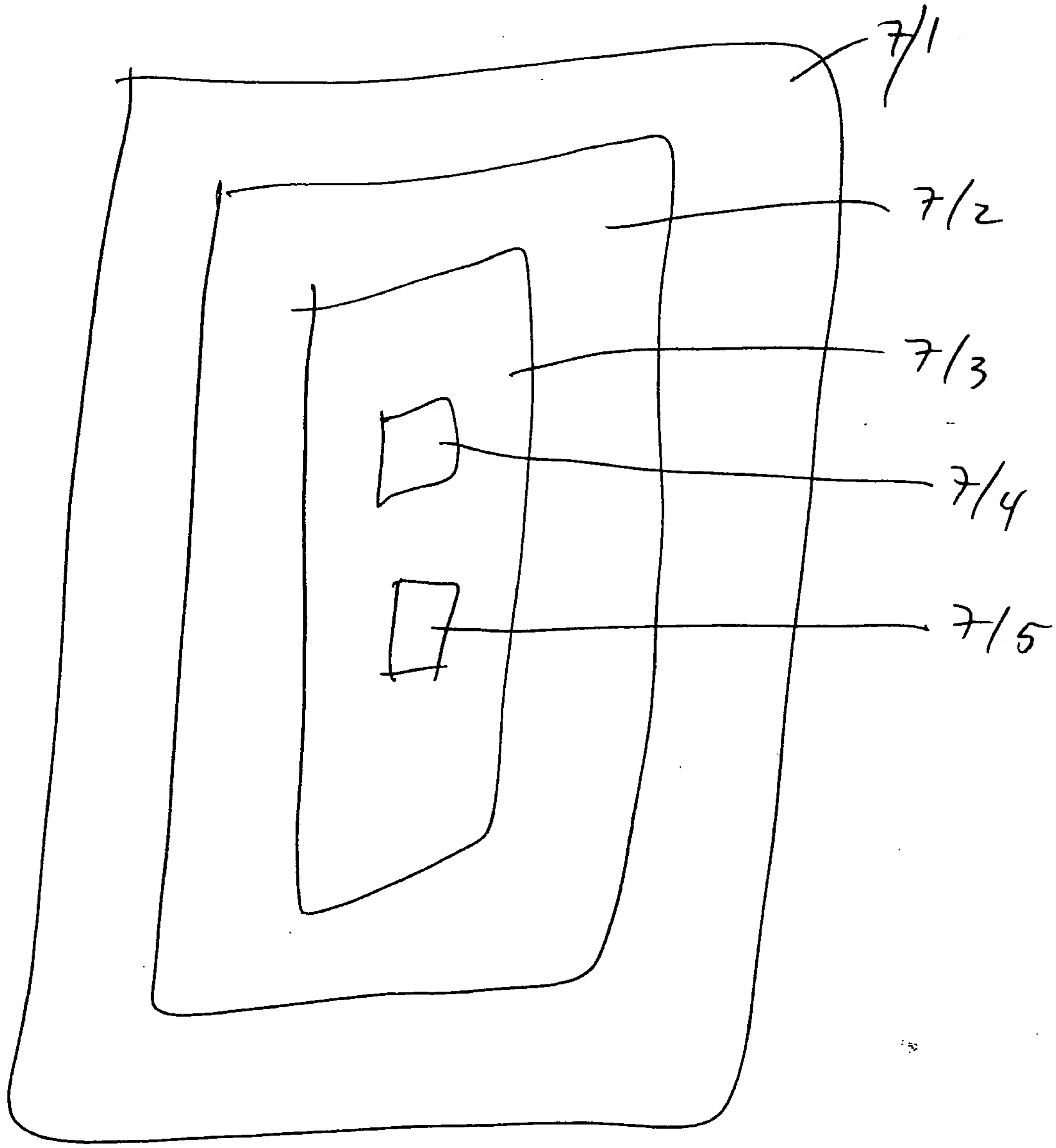


FIG 7

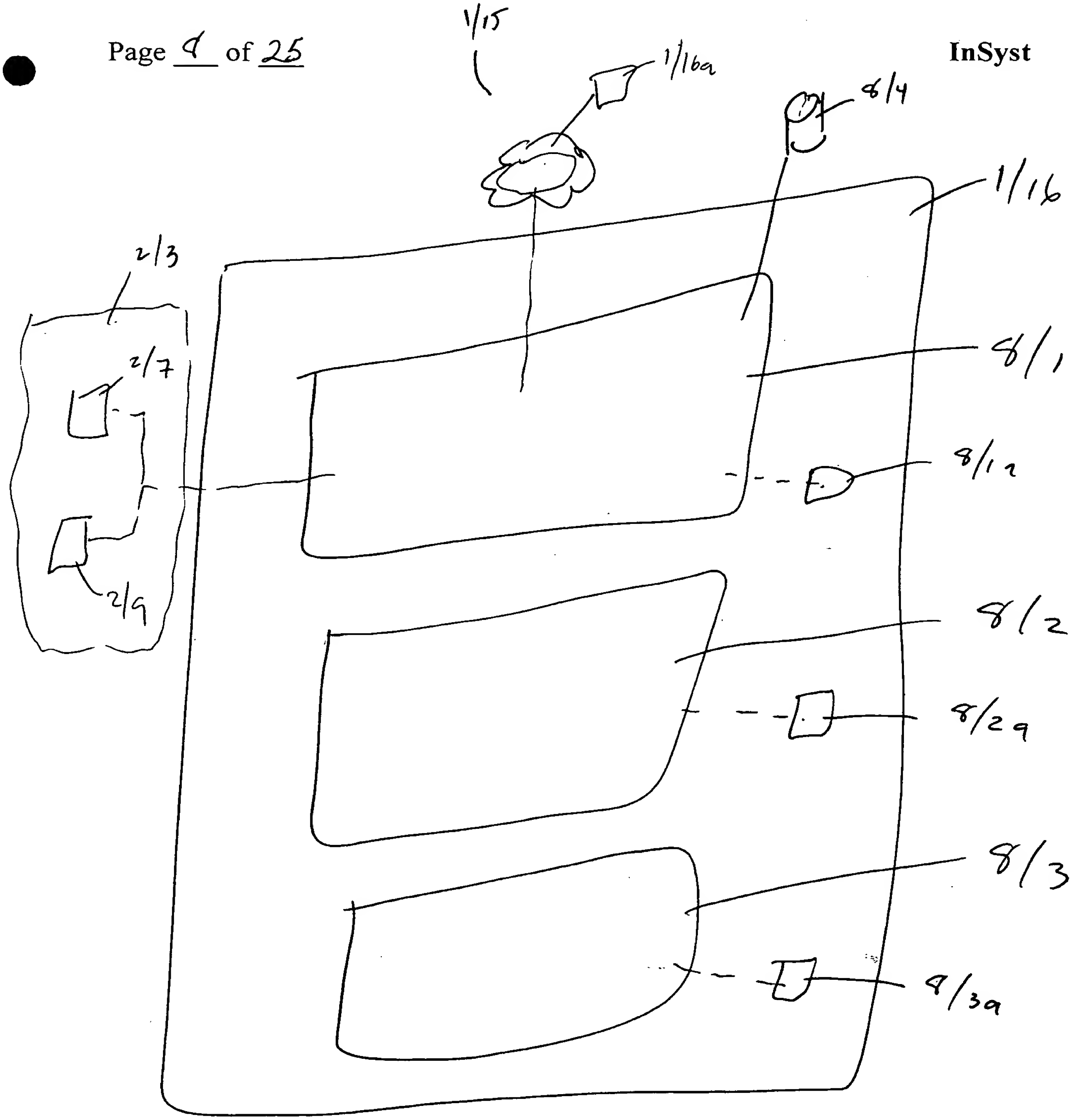


FIG 8

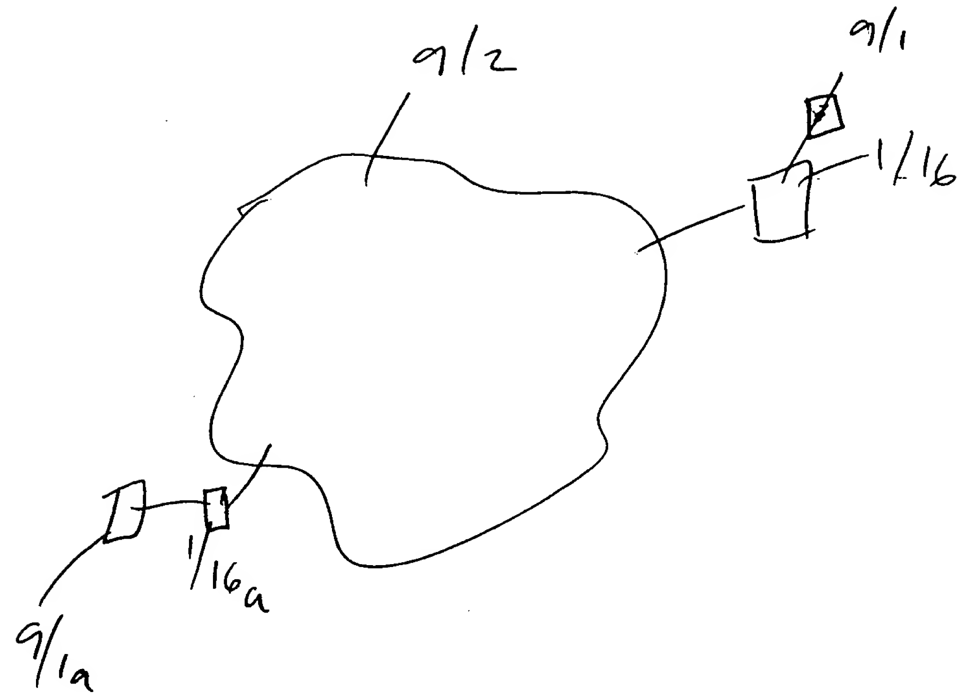


FIG 9

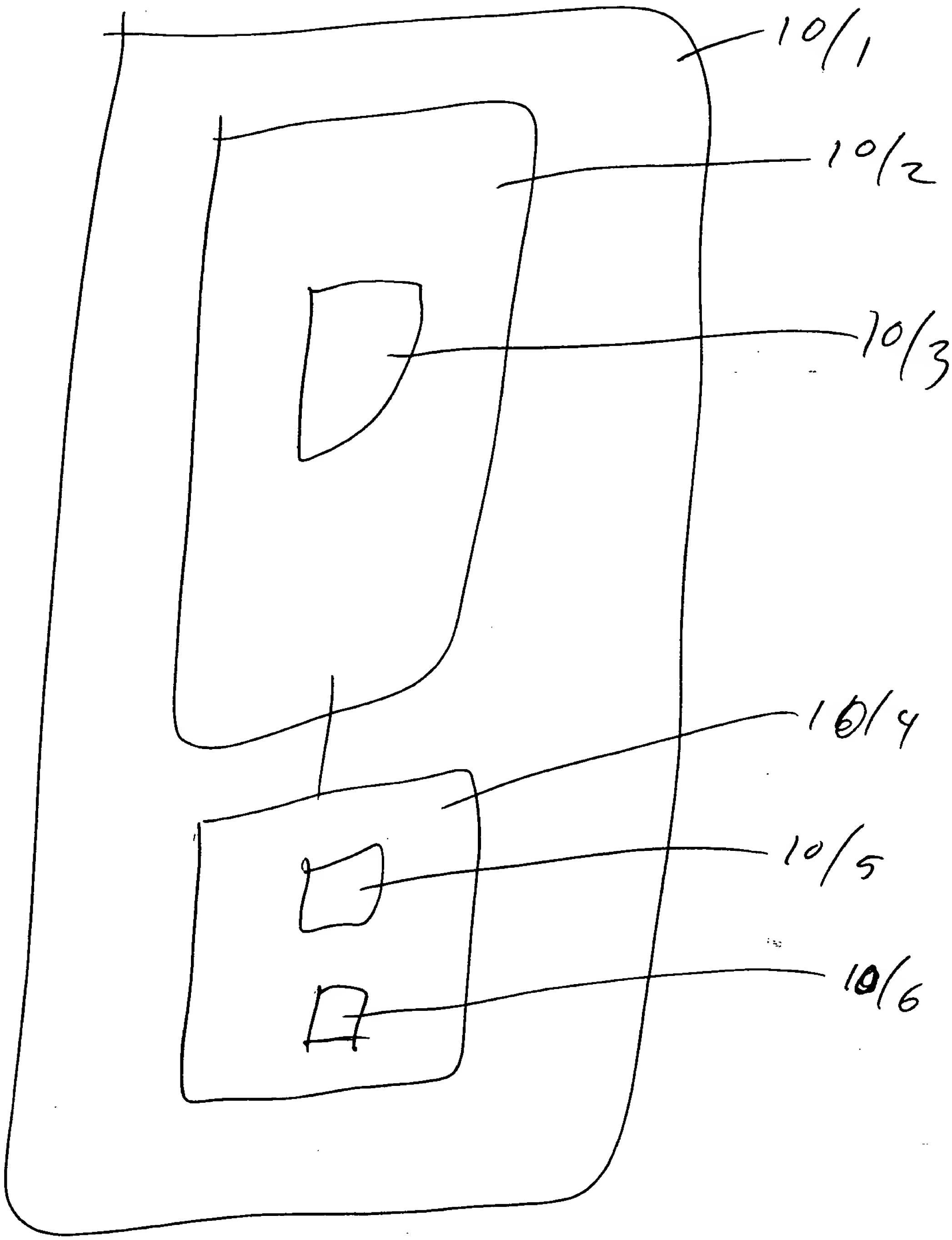


FIG 10

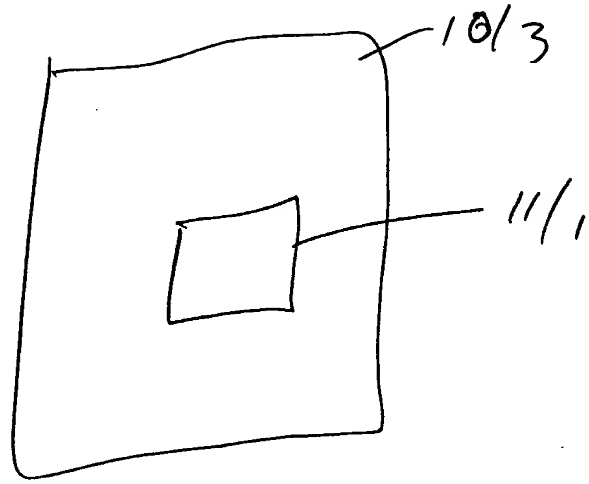


FIG 11

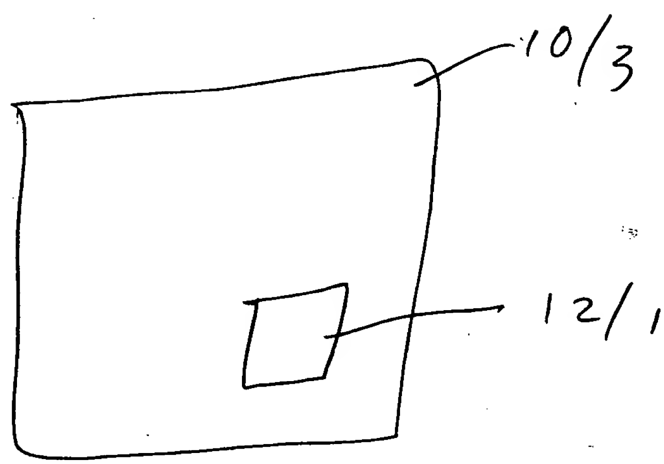


FIG 12

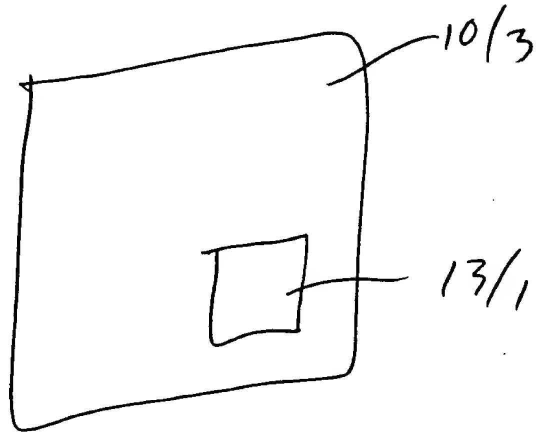


FIG 13

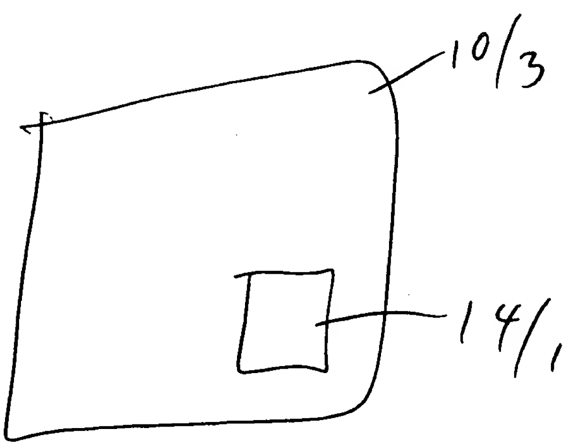


FIG 14

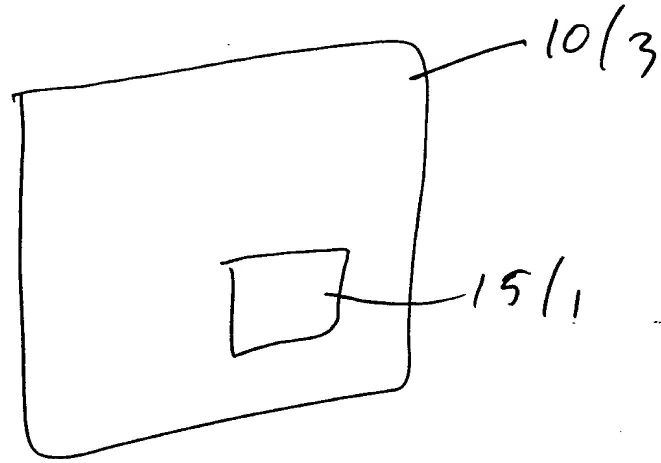


FIG 15

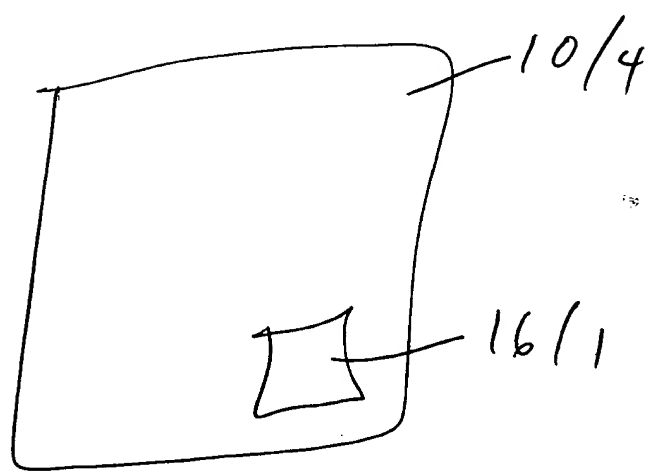


FIG 16

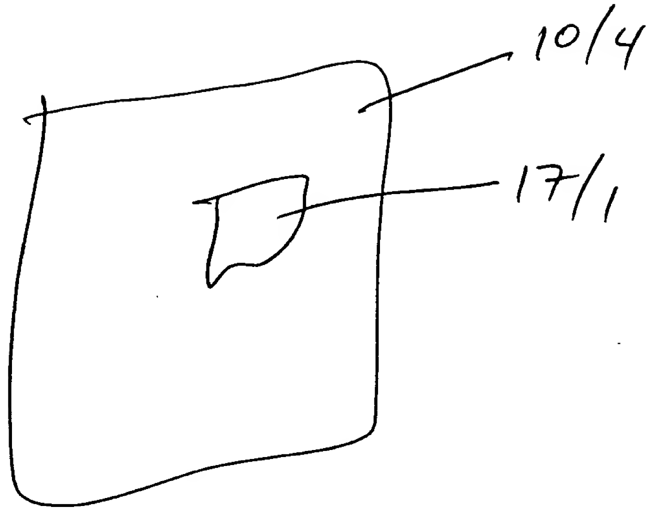


FIG 17

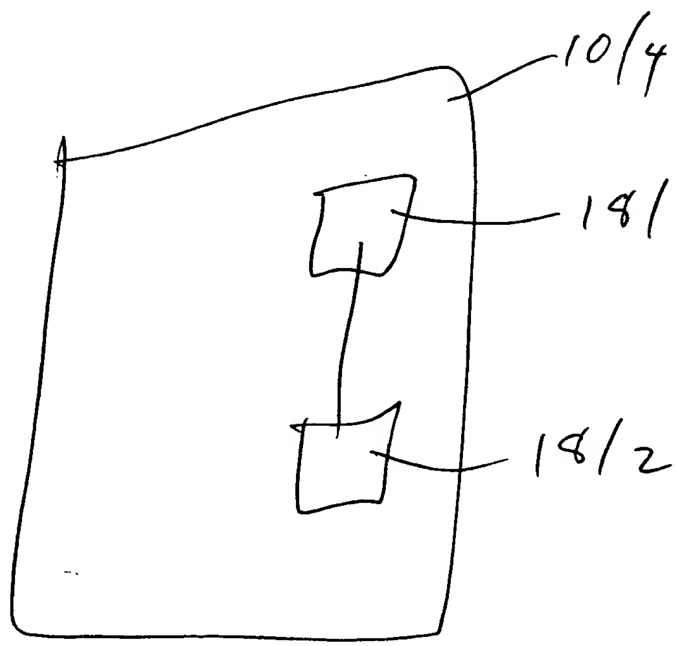


FIG 18

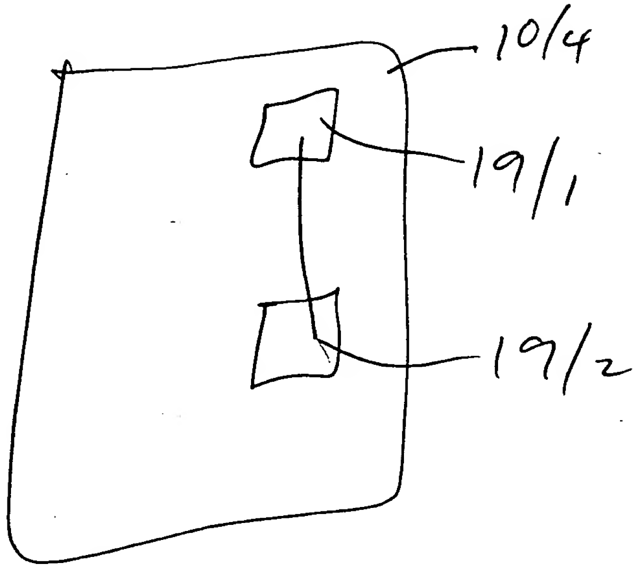


FIG 19

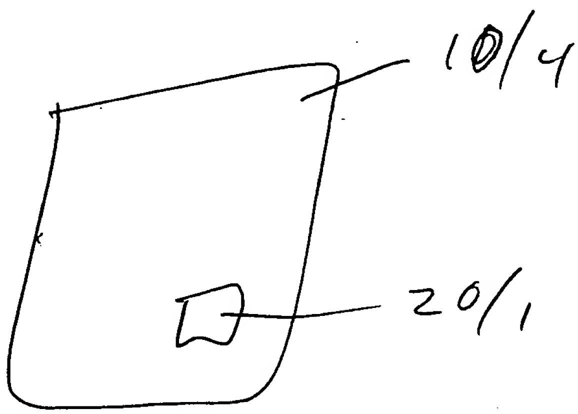


FIG 20

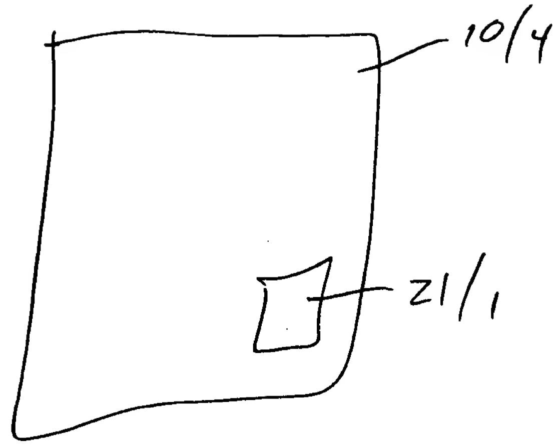


FIG 21

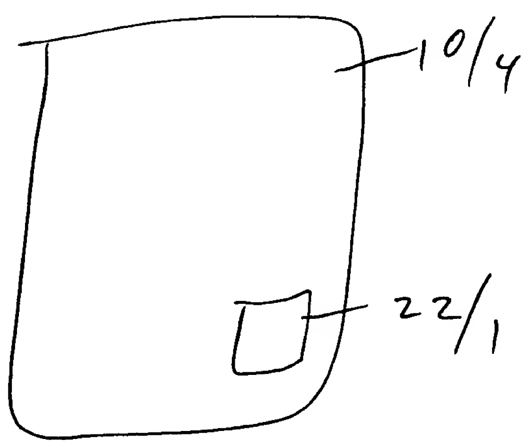


FIG 22

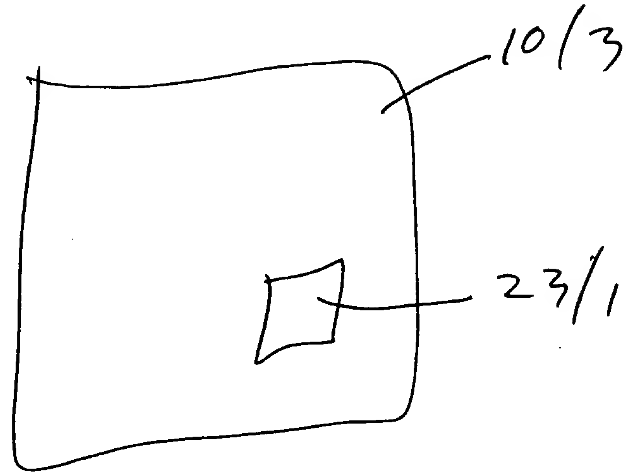


FIG 23

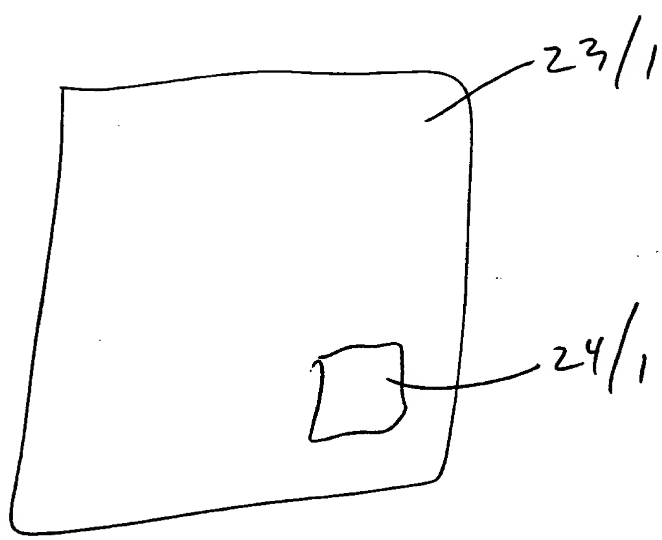
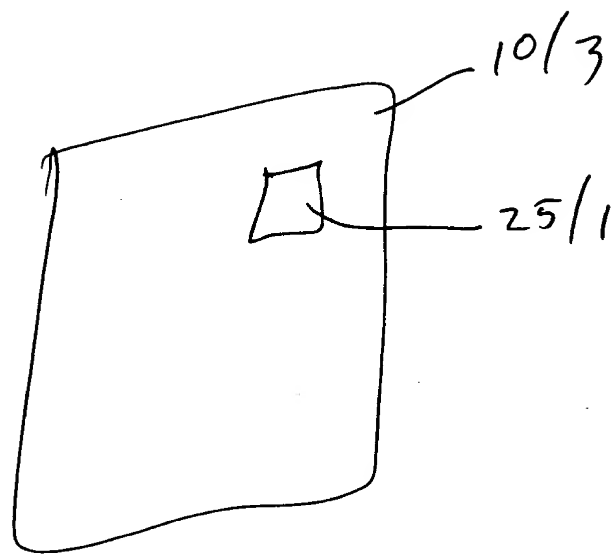
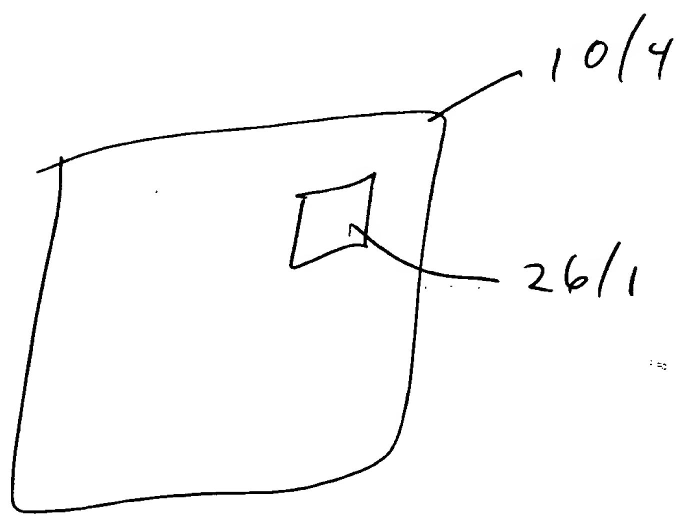


FIG 24



F16 25



F16 26

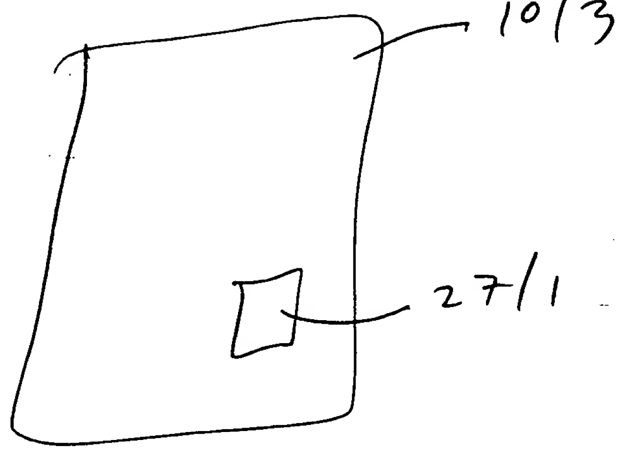


FIG 27

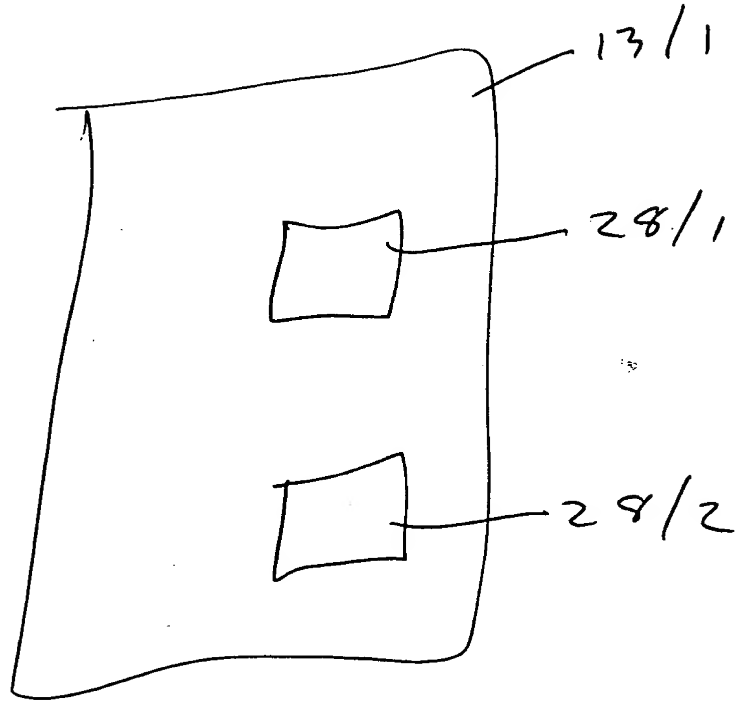


FIG 28

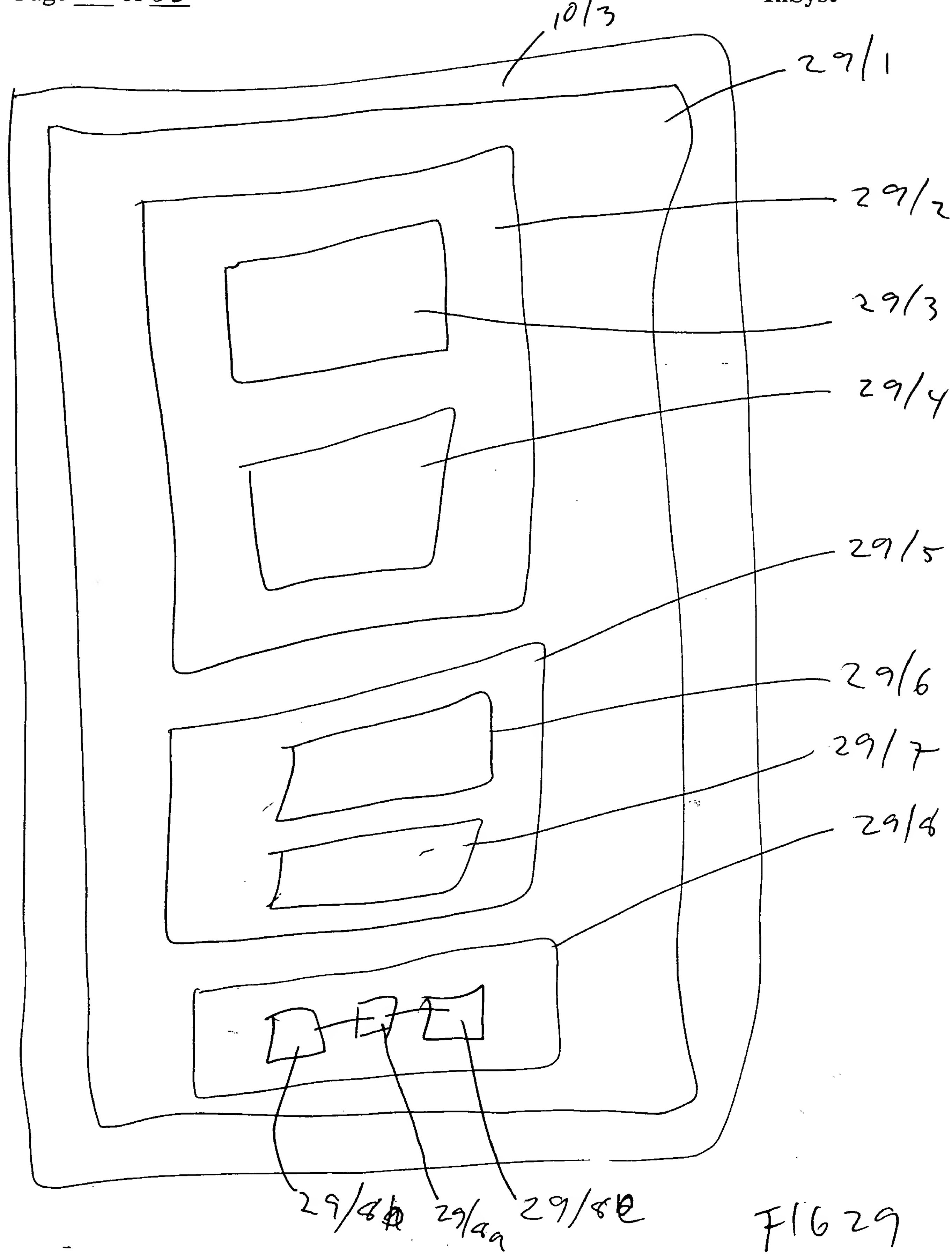


FIG 29

ILLUSTRATION I

InSyst

Diagram 2.4 The Mapping of the Photolithography Process

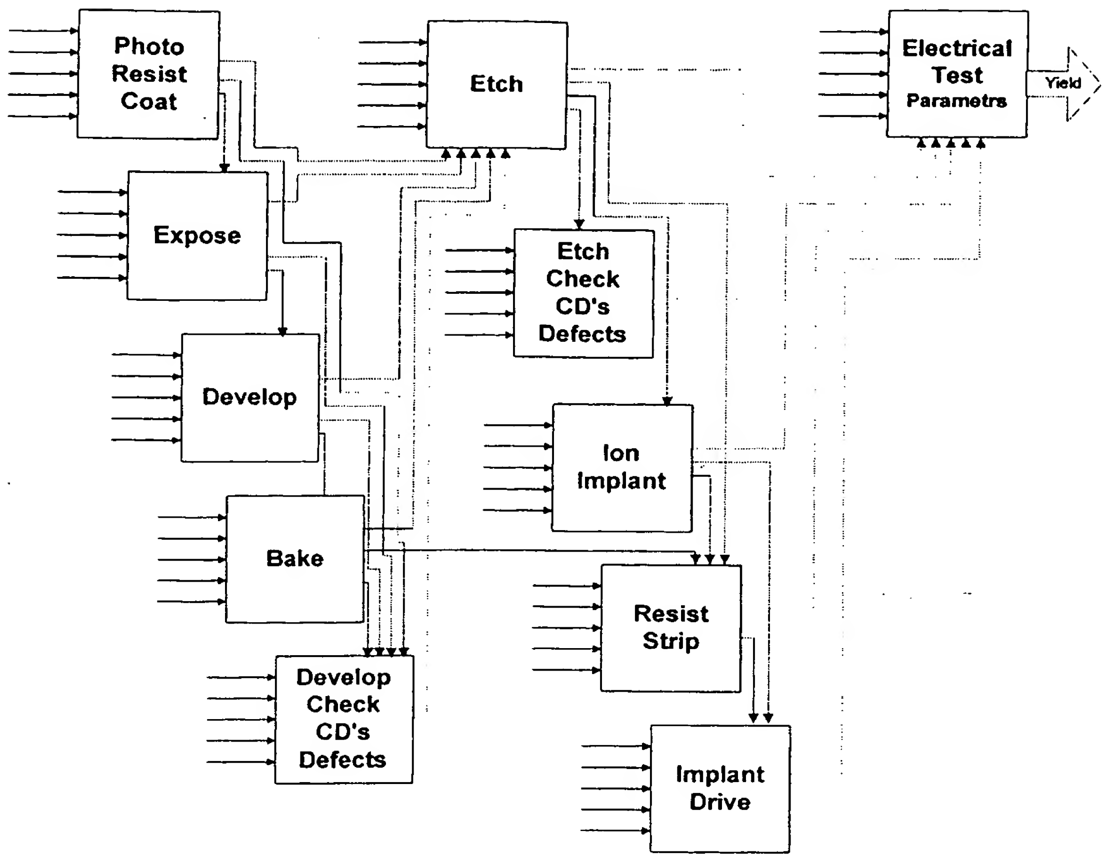
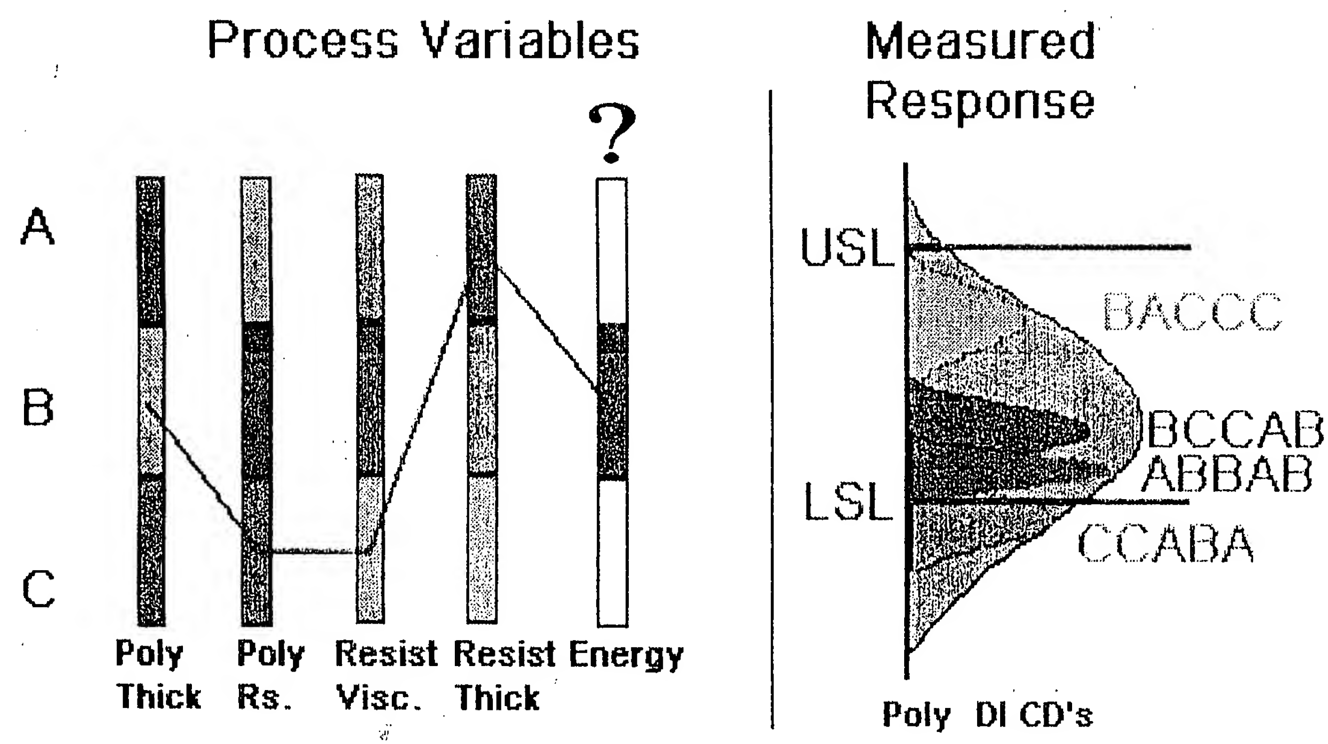
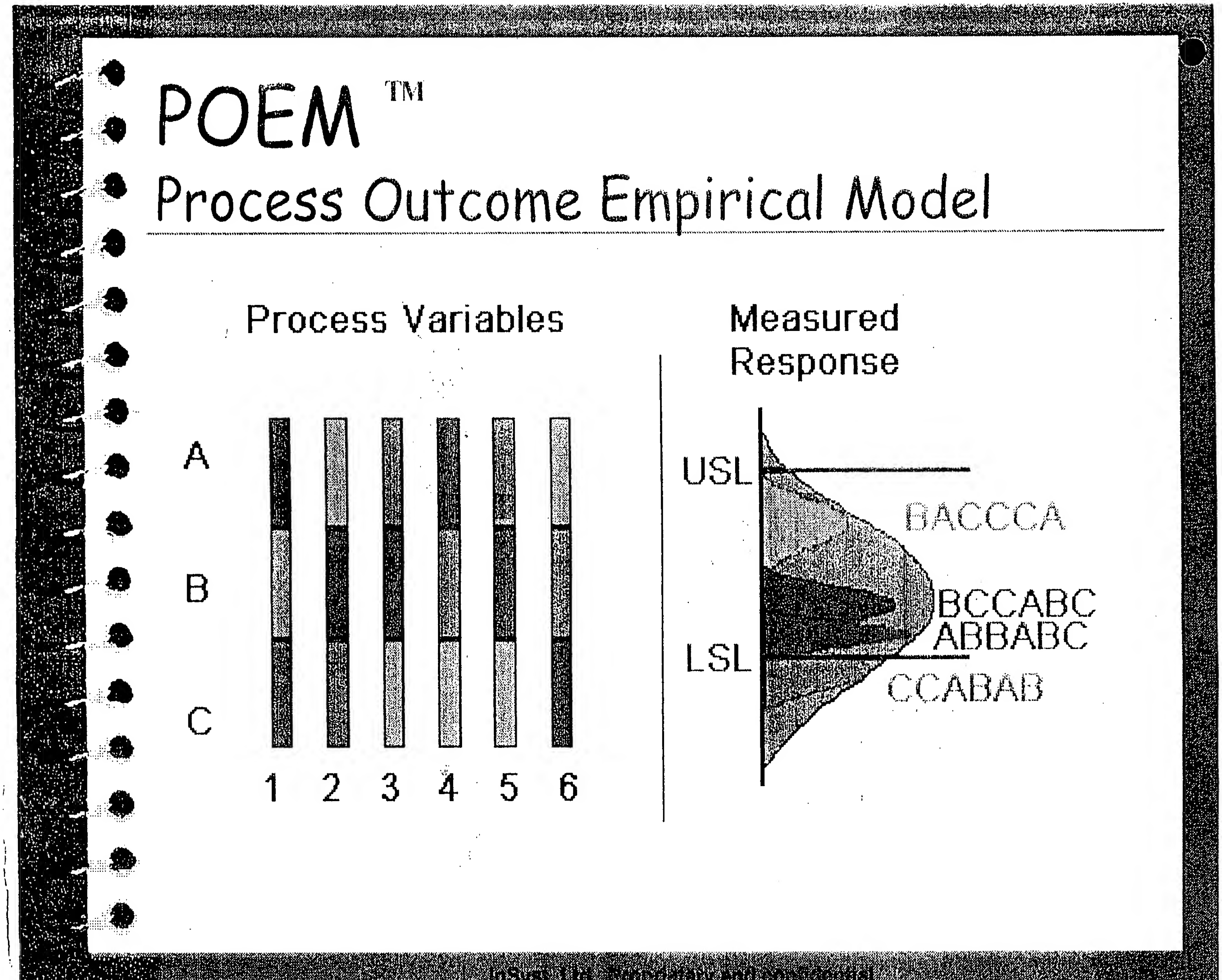


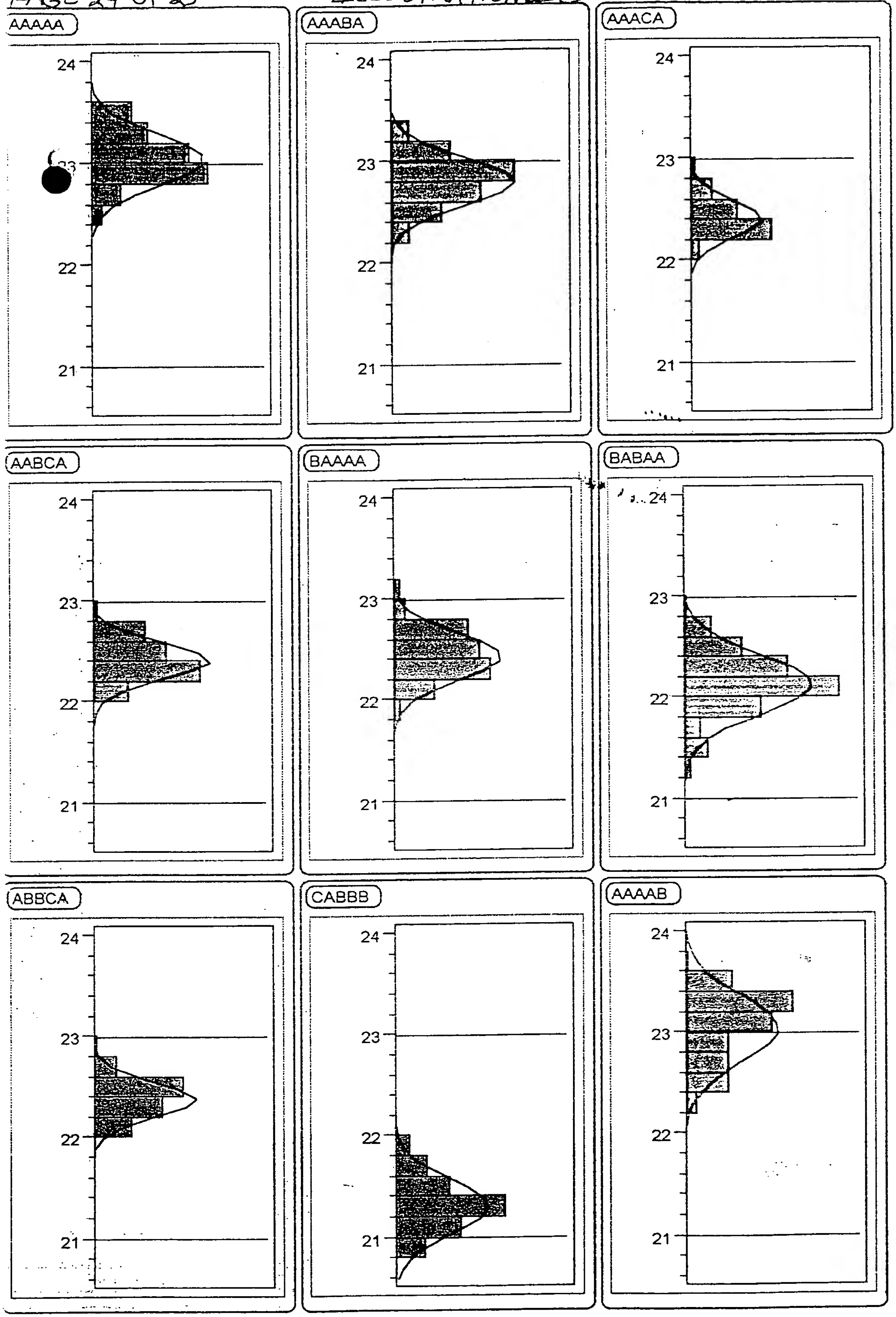
ILLUSTRATION 2A

InSyst

Feed Forward Optimization

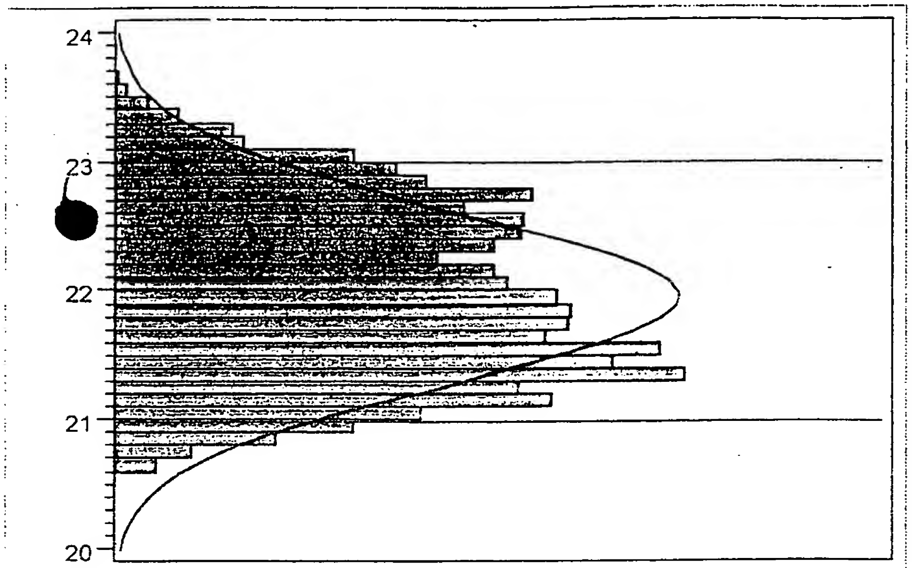



POEM ™
Process Outcome Empirical Model
Process Variables
A
B
C
1 2 3 4 5 6
Measured Response
USL
LSL
BACCCA
BCCABC
ABBABC
CCABAB
InSyst Ltd. Proprietary and confidential



InSyst Ltd. Proprietary and confidential

output



Quantiles

maximum	100.0%	23.830
	99.5%	23.412
	97.5%	23.210
	90.0%	22.890
quartile	75.0%	22.520
median	50.0%	21.920
quartile	25.0%	21.430
	10.0%	21.130
	2.5%	20.870
	0.5%	20.710
minimum	0.0%	20.590

Moments

Mean	21.97
Std Dev	0.66
Std Error Mean	0.01
Upper 95% Mean	21.98
Lower 95% Mean	21.96
N	11968.00
Sum Weights	11968.00

Capability Analysis

Specification	Value	Percent	Actual	Normal
Lower Spec Limit	21	%Below LSL	5.381	7.095
Upper Spec Limit	23	%Above USL	6.451	5.959
Spec Target	22			
Sigma	0.660716			

Capability	Index
CPL	0.490
CPU	0.519
CPK	0.490
CP	0.505
CPM	0.504